

738 N

Founded by G. Stanley Hall in 1891, reorganized by Carl Murchison in 1925

THE PEDAGOGICAL SEMINARY AND

JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior,
and Comparative Psychology

EDITED BY
CARL MURCHISON

JOHN E. ANDERSON
University of Minnesota

CHARLOTTE BÜHLER
University of
Southern California

CYRIL BURT
University of London

LEONARD CARMICHAEL
Smithsonian Institution

SANTE DE SANCTIS
R. Università di Roma

ARNOLD GESELL
Yale University

WILLIAM HEALY
Judge Baker Foundation,
Boston

Arthur T. Jersild
Teachers College,
Columbia University

BUFORD JOHNSON
The Johns Hopkins
University

HAROLD E. JONES
University of California

TRUMAN L. KELLEY
Harvard University

K. S. LASHLEY
Harvard University

HENRI PIÉRON
Université de Paris

Robert R. Sears
Harvard University

GEORGE D. STODDARD
University of Illinois

CALVIN P. STONE
Stanford University

LEWIS M. TERMAN
Stanford University

GODFREY THOMSON
University of Edinburgh

C. J. WARDEN
Columbia University

JOHN B. WATSON
New York City

If this space should be unstamped, this is the regular library edition.
But if this space is stamped with a designating title, this is a special
edition, sold under the restrictions of a bilateral contract, and may not
be resold for a period of five years from date of publication.

VOLUME 82
1953

Copyright, 1953, by The Journal Press
Published quarterly by The Journal Press
Provincetown, Massachusetts, U. S. A.



AUTHOR INDEX

Allen, Robert M.	319
Ames, Louise B.	183
Ammons, Carol H.	205
Ammons, Robert B.	205
Brown, W. Lynn	11
Darcy, Natalie T.	21
Fredericson, Emil	143
Frenkel-Brunswik, Else	91
Garza, Cesar Octavio de la	11
Greenhut, Ann	155
Havel, Joan	91
Jackson, Joseph	215
Johnson, Granville B., Jr.	3
Kline, Milton V.	137
Learned, Janet	183
Metraux, Ruth	183
Norman, Ralph	69
Norton, Joseph L.	235, 263
Ross, Sherman	279, 299, 309
Schneidman, Edwin S.	149
Sievers, Dorothy J.	69
Smith, Wendell	279, 299, 309
Walker, Richard	183
Wilson, Frank T.	59
Young, Francis A.	155

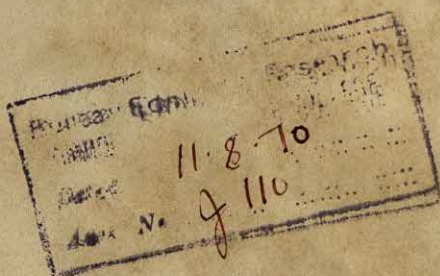


TABLE OF CONTENTS

Bilingualism as measured by a reaction-time technique and the relationship between a language and a non-language intelligence quotient	3
BY GRANVILLE B. JOHNSON, JR.	
The effect of electroconvulsive shock on the anticipatory responses of the rat in a semi-linear maze	11
BY W. LYNN BROWN AND CESAR OCTAVIO DE LA GARZA	
A review of the literature on the effects of bilingualism upon the measurement of intelligence	21
BY NATALIE T. DARCY	
Some special ability test scores of gifted children	59
BY FRANK T. WILSON	
Some suggestive results in psychometric testing of the cerebral palsied with Gesell, Binet, and Wechsler scales	69
BY DOROTHY J. SIEVERS AND RALPH D. NORMAN	
Prejudice in the interviews of children: I. Attitudes toward minority groups	91
BY ELSE FRENKEL-BRUNSWIK AND JOAN HAVEL	
SHORT ARTICLES AND NOTES	137
Childhood fears in relation to hypnotic age regression: A case report (Milton V. Kline)	137
The wall-seeking tendency in three inbred mouse strains (<i>Mus musculus</i>) (Emil Fredericson)	143
BOOKS	147
CRITICAL REVIEWS OF RECENT BOOKS	149
Bühler, C., Smutter, F., Richardson, S., & Bradshaw, F. <i>Childhood Problems and the Teacher</i> . (Reviewed by Edwin S. Shneidman)	149
BOOKS RECENTLY RECEIVED	151
Visual depth perception in the rat	155
BY ANN GREENHUT AND FRANCIS A. YOUNG	
Development of perception in the young child as observed in responses to the Rorschach test blots	183
BY LOUISE B. AMES, JANET LEARNED, RUTH METRAUX, AND RICHARD WALKER	
Aggression in doll-play: Interviews of two- to six-year-old white males	205
BY CAROL H. AMMONS AND ROBERT B. AMMONS	
A statistical analysis of an alumni survey	215
BY JOSEPH JACKSON	
Patterns of vocational interest development and actual job choice	235
BY JOSEPH L. NORTON	
General motives and influences in vocational development	263
BY JOSEPH L. NORTON	

The hoarding behavior of the mouse: I. The rôle of previous feeding experience	279
BY WENDELL SMITH AND SHERMAN ROSS	
The hoarding behavior of the mouse: II. The rôle of deprivation, satiation, and stress	299
BY SHERMAN ROSS AND WENDELL SMITH	
The hoarding behavior of the mouse: III. The storing of "non-relevant" material	309
BY WENDELL SMITH AND SHERMAN ROSS	
BOOKS	317
CRITICAL REVIEWS OF RECENT BOOKS	319
Greene, E. B. <i>Measurements of Human Behavior</i> . (Reviewed by Robert M. Allen)	319
BOOKS RECENTLY RECEIVED	323

BILINGUALISM AS MEASURED BY A REACTION-TIME
TECHNIQUE AND THE RELATIONSHIP BETWEEN
A LANGUAGE AND A NON-LANGUAGE
INTELLIGENCE QUOTIENT*

Division of Teacher Education, Emory University, Georgia

GRANVILLE B. JOHNSON, JR.

A. THE PROBLEM

Administration of objective tests of intelligence to individuals possessing a language other than English as the primary language often yields invalid results. This may not be so much the result of the testee's inability to read, but rather his lack of acquaintance with the nuances of culture which compose, in part, the intelligence test.

A proposed method of ascertaining the extent to which the dominant culture has been assimilated by the bilingual is to measure the knowledge of the language of that culture. It may be assumed then that knowledge of the language is reflective of the degree to which the individual has assimilated the cultural elements of which the language is representative.

With reference to this problem the present study was developed. The conclusions pertain only from data secured from 30 Spanish boys from a specific town in the southwest United States.

B. PURPOSE

The purpose of the present study was to ascertain the relationship existing between a language and a non-language test of intelligence, and to note the rôle which the degree of bilingualism as measured by a reaction-time technique plays in this relationship.

C. REVIEW OF PREVIOUS STUDIES

Means of measurement of bilingualism may be classified into three categories: (a) Background questionnaires, (b) Association techniques, and (c) Rating scales (1).

The techniques employed in the present investigation utilized the "Hoffman Bilingual Schedule" together with the "Reaction-Time Technique."

The relationship between bilingualism and intelligence test scores has

*Received in the Editorial Office on December 25, 1950.

interested several investigators. In general, the technique employed was to measure the bilingual background of the group of subjects under consideration, and then to find the relationship between this score and scores on intelligence or other performance tests (2, 3, 4, 5, 7, 8, 9, 10, 12). The results of these studies revealed that increase in English is associated with increase in mental age. Pintner compared foreign language groups with English speaking groups for achievement on verbal tests as against achievement on non-verbal tests of intelligence.

His findings demonstrated that the foreign language groups performed better on non-language tests than on verbal tests of intelligence. He attributes the inferior rating on the verbal tests to language handicap (14, 15, 16, 17).

D. SUBJECTS AND MATERIALS

1. *Subjects*

Thirty Spanish boys from the southwest region of the United States were the subjects of this investigation. They were further delimited in that their ages ranged from 9 through 12 years and were in school grades four through six. All of the boys used English in school and each had a knowledge of the Spanish language.

2. *Materials*

1. The Otis Self-Administering Test of Mental Ability (Intermediate (13)).

2. The Goodenough Draw a Man Test (6).

3. The Hoffman Test of Bilingualism (11). This is a test of bilingual background developed to furnish a quantitative measure of the bilingual background to which the individual has been exposed. The validity of this measure is $.73 \pm .03$; the reliability is .85.

4. The Reaction-Time Test of Bilingualism. This is an instrument devised to measure the facility with which an individual can manipulate the symbols of language and to compare this facility for each of the two languages so that the degree of bilingualism may be measured. Since the study of language may be reflective of the degree of cultural assimilation, the test first measures the number of English (Spanish) words an individual can recall in five minutes and then compares this with the number of Spanish (English) words he can recall in the same length of time.

Thus a ratio may be derived which may then be correlated with discrepancy between the results on the language and the non-language tests of intelligence.

The administration of the test requires a room as far as possible devoid of extrinsic visual or auditory stimuli in which the subject sits facing a blank wall. After rapport has been established, the administrator says: "I want you to name all the words in English (Spanish) which you can think of. Any words will be all right. Name as many words as you can; I will tell you when to stop." The investigator, who is of a stereotype associated with the language he is testing, tabulates each word in one minute divisions for five minutes (the total for five minutes was used in the statistical analysis). At least two weeks after this time (to relieve practice effect) another administrator (Spanish for the Spanish section of the test and Anglo for the English section) tests in the same way the other language under consideration. In the present investigation, half the subjects were tested first for English words and the other half for Spanish words first. The difference between the two sequences was not found to be significant.

E. METHOD OF PROCEDURE

The Goodenough test of intelligence was administered to each of the subjects and was followed by the Otis. It was found that the establishment of rapport was facilitated by presentation of the Goodenough first.

After the tests of intelligence had been administered, the team of three persons administered the tests of bilingualism.

Relationships between the Otis and Goodenough, Hoffman and Goodenough, Otis and Reaction-Time, and Goodenough and Reaction-Time were established.

The Otis and the Goodenough results were compared and a ratio between them was derived which was then related with the Hoffman and then with the Reaction-Time Test.

The degree of relationship between the Hoffman and the Reaction-Time Test was ascertained.

F. RESULTS

Table 1 is a presentation of the significant data derived from the 30

TABLE 1
MEAN AND STANDARD DEVIATION OF THE FOUR MEASURES FOR THE THIRTY SPANISH SUBJECTS

	M	σM	σ	$\sigma\sigma$
Goodenough (<i>IQ</i>)	98.77	1.80	9.72	1.28
Otis (<i>IQ</i>)	86.37	1.63	8.76	1.15
Hoffman (score)	1.90	.12	.64	.08
Reaction-time (Span.)	53.77	3.33	17.91	2.35
Reaction-time (Eng.) (words per five minutes)	71.13	4.14	22.29	2.93

Spanish subjects which compose the population under consideration. Table 1 demonstrates that the Goodenough intelligence quotient for the Spanish subjects was almost average for the total population while the Otis intelligence quotient was considerably below average. The difference between these two was significant at the .01 level ($t = 5.10$).

The Reaction-Time Test of Bilingualism scores reveal that the subjects were able to respond with more English words than Spanish at an approximate ratio of seven to five.

The relationships existing between the intelligence tests and the bilingual tests were established. Table 2 presents these relationships.

Analysis of Table 2 reveals some rather interesting results. The Hoff-

TABLE 2
THE RELATIONSHIPS BETWEEN THE INTELLIGENCE TESTS AND THE TESTS OF BILINGUALISM

	Hoffman		Reaction-time	
	r	σr	r	σr
Otis	-.121	.183	-.412	.154
Goodenough	.054	.185	.425	.152

man and the Reaction-Time Tests appear to be measures of common factors since their relationships with the tests of intelligence, though varying quantitatively, are in the same direction. The Reaction-Time Test appears to be a more sensitive measuring instrument than the Hoffman in this particular situation because where the Hoffman relationships manifest insignificant values, the Reaction-Time relationships are at the .02 level of significance.

The negative relationships between the Otis and the tests of bilingualism indicate that there is a negative relationship between performance on the Otis and degree of bilingualism. Specifically, with reference to the Reaction-Time Test, it may be concluded that there is, at the .02 level of significance, an inverse relationship between performance on the Otis and knowledge of Spanish in comparison with English. Or, in conclusion, the higher the Otis *IQ* the less knowledge of Spanish in comparison with English.

The reverse of the Table 2 finding was made when the Goodenough was compared with the measures of bilingualism. That is: the degree of bilingualism is associated with superior performance on the Goodenough. Specifically, with reference to the Reaction-Time Test, a tendency was revealed which indicated that greater knowledge of Spanish in comparison with English is directly associated with Goodenough performance.

The relationship between the two measures of intelligence was found to be insignificant ($r = .022$). Evidently they are tests of separate factors.

The relationships existing between the Goodenough-Otis discrepancy and the Reaction-Time Test; the Goodenough-Otis discrepancy and the Hoffman; and the Hoffman and the Reaction-Time Test are presented in Table 3.

Table 3 shows one significant correlation: that between the Goodenough-

TABLE 3
RELATIONSHIPS BETWEEN INTELLIGENCE TEST RATIO AND BILINGUAL TESTS

	r	σr	Significance
Goodenough-Otis Discrepancy r Reaction-Time Test	.633	.111	.01 ($t = 4.33$)
Goodenough-Otis Discrepancy r Hoffman Test	.044	.182	Insignificant
Hoffman r Reaction-Time Test	.116	.156	Insignificant

Otis discrepancy and the Reaction-Time Test of Bilingualism. This indicates that the greater knowledge of English in comparison with knowledge of Spanish is associated with less discrepancy between the Otis and Goodenough *IQ* scores. The Otis results most nearly approach Goodenough results in those individuals with fewest Spanish responses on the Reaction-Time Test. Although only 13 per cent of the Reaction-Time Test, the remaining subjects showed less discrepancy between Goodenough and Otis *IQ*. Scores in direct proportion to the degree of ascendance of English words to Spanish.

G. SUMMARY AND CONCLUSIONS

An intelligence test employing the English language is probably not a valid measuring instrument when employed with subjects deficient in the assimilation of the culture of which English is reflective.

To ascertain the degree of cultural assimilation for each subject, the Reaction-Time Test of Bilingualism was developed. Though the technique of naming as many words as possible in a given length of time has been used as a measure of intelligence (e.g., Binet), the factor of intelligence is ruled out in the Reaction-Time Test because the score of bilingualism on this test is a ratio between number of words given in English and the number of words given in Spanish. This ratio was correlated with another test of bilingualism, the Hoffman Test of Bilingualism, and it was found that little relationship existed between the two.

Coefficients of correlation between scores on the tests of intelligence and scores on the tests of bilingualism revealed that the Hoffman and the Reac-

tion-Time Tests appear to be measures of common factors since their relationships with the intelligence tests, though varying quantitatively, are in the same direction. The Reaction-Time Test appears to be the more sensitive measuring instrument of the two.

There was a negative relationship between performance on the Otis and the degree of bilingualism. The relationship between the tests of bilingualism and the Goodenough, however, indicated that the degree of bilingualism is associated with superior response on this performance test of intelligence.

The relationship between the two measures of intelligence was found to be insignificant ($r = -.022$), revealing that they are probably tests of separate factors.

A coefficient of correlation was derived between the Reaction-Time Test of Bilingualism and the discrepancy between the Otis Test of intelligence and the Goodenough "Draw a Man Test." The relationship was found in this case to be significant at the .01 level ($r = .663$); the Hoffman and the Goodenough-Otis discrepancy relationship, however, was insignificant as was the relationship between the Hoffman and the Reaction-Time Test.

It was found that greater knowledge of English in comparison with knowledge of Spanish is associated with less discrepancy between the Otis and the Goodenough intelligence quotient scores. The Otis results most nearly approached Goodenough results in those individuals with fewest Spanish responses on the Reaction-Time Test. Although only 13 per cent of the subjects yielded a greater number of Spanish words than English on the Reaction-Time Test, the remaining subjects showed less discrepancy between Goodenough and Otis intelligence quotient scores in direct proportion to the degree of ascendance of English words to Spanish.

It may be concluded from the findings of this study that measuring the intelligence of bilingual subjects presents complex problems which possibly render both linguistic (Otis) and performance (Goodenough) tests invalid.

REFERENCES

1. ARSENIAN, S. Bilingualism in the postwar world. *Psychol. Bull.*, 1945, **42**, 71.
2. ANDREWS, F. E. The handicap of a limited vocabulary. *Chicago Sch. J.*, 1928, **11**, 128-130.
3. BERE, M. A Study of Comparative Intelligence of Children of Foreign Parentage. Contributions to Education, No. 154. New York: Bureau of Publications, Teachers College, Columbia University, 1924. Pp. 105.
4. BOVET, P. Bilinguisme et Education. Report of the 1932 Conference on Bilinguisme at Nice. Pp. 14.
5. CONFERENCE INTERNATIONALE SUR LE BILINGUISME. Luxembourg, 1928. Le Bilinguisme et l'Education. Travaux de la conference internationale tenue a Luxembourg. Geneve: Bureau International d'Education. Pp. 184.

6. GOODENOUGH, F. L. Measurement of Intelligence by Drawings. New York: World Book, 1926. Pp. 177.
7. HAUGHT, B. F. The language difficulty of Spanish-American children. *J. Appl. Psychol.*, 1931, **15**, 92-52.
8. HOFFMAN, N. N. H. The Measurement of Bilingual Background. Contributions to Education, No. 623. New York: Bureau of Publications, Teachers College, Columbia University, 1934. Pp. 75.
9. LADD, M. R. The Relation of Social, Economic, and Personality Characteristics to Reading Ability. Contributions to Education, No. 582. New York: Bureau of Publications, Teachers College, Columbia University, 1933. Pp. 100.
10. MADSEN, I. N. Some results with the Stanford revision of the Binet-Simon tests. *Sch. & Soc.*, 1924, **19**, 559-562.
11. MEAD, M. Group intelligence tests and linguistic disability among Italian children. *Sch. & Soc.*, 1927, **24**, 465-468.
12. MURDOCH, K. D., MADDOW, D., & BERG, N. L. A study of the relation between intelligence and the acquisition of English. *Nature and Nurture*, Twenty-Seventh Yearbook of the National Society for the Study of Education. Bloomington: Public School Publishing Company, 1928.
13. OTIS, A. S. Otis Self-Administering Test of Mental Ability. Intermediate Form. New York: World Book.
14. PINTNER, R. Comparison of American and foreign children on intelligence tests. *J. Ed. Psych.*, 1923, **14**, 292-295.
15. ———. Results obtained with the non-language group test. *J. Ed. Psych.*, 1924, **15**, 473-483.
16. ———. The influence of language background on intelligence tests. *J. Ed. Psych.*, 1932, **3**, 235-240.
17. PINTNER, R., & KELLER, R. Intelligence tests of foreign children. *J. Ed. Psych.*, 1922, **13**, 214-222.

Division of Teacher Education

Emory University

Emory University, Georgia

THE EFFECT OF ELECTROCONVULSIVE SHOCK ON THE ANTICIPATORY RESPONSES OF THE RAT IN A SEMI-LINEAR MAZE*

Department of Psychology, University of Texas

W. LYNN BROWN¹ AND CESAR OCTAVIO DE LA GARZA

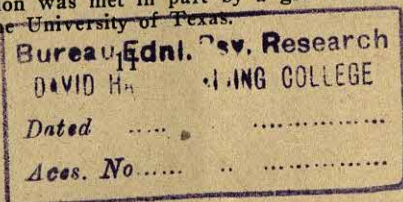
A. PROBLEM AND LITERATURE

The present experiment was designed to study the effect of electroconvulsive shock on the anticipatory responses of the rat in a semi-linear maze. Extensive study has been made in recent years on the effect of electroconvulsive shock on learning and retention of maze habits. In general, these studies indicate a decrement in maze performance following electroconvulsive shock. However, Russell (10) found that very simple habits do not appear to be impaired following electroconvulsive shocks.

Fosmire and Brown (5), using a maze task which "required high level spatial organization in the white rat," investigated the effect of electroconvulsive shock on spatial learning. They found that after the animals had mastered their task to a criterion of four errorless runs, subsequent daily post-run shocks produced an error increase. Further investigation on this problem by Brown and Fosmire (3) confirmed their belief that the results obtained were a function of the electroconvulsive shock. Brown and Wilbanks (4) attempted to isolate the number of shocks beyond which no further impairment in learning is apparent on the spatial problem studied by Fosmire and Brown (5) and Brown and Fosmire (3). One group of animals (Group II) was subjected to one daily electroconvulsive shock after each trial for the first six days of the experiment. A second group (Group III) was subjected to one daily electroconvulsive shock for the first 12 days of the experiment, after which no further shocks were administered. A third group (Group IV) was shocked daily for the first 18 days of the experiment. All groups ran the maze for 30 days. Brown and Wilbanks (4) found that "although the 6-day and the 12-day shock groups were able to reduce their errors, they never reached the level of efficiency exhibited by the control group. The animals in the 18-day shock group never per-

*Received in the Editorial Office on February 14, 1951.

¹The cost of this investigation was met in part by a grant to the senior author by the Research Institute of the University of Texas.



formed better than slightly under chance." Brown and Wilbanks suggested that the reason their animals failed to learn the spatial problem may have been "due to the fact that electroshock convulsions prevented the formation of a stable memory trace." Brown and Wilbanks observed a tendency for shocked animals to arrive at "a crude solution to the problem in the form of position habits" when they could not learn the spatial problem. They concluded that "electroshock convulsions evidently reduced spatial learning to response learning."

Anticipation as a source of error in the learning of serial responses has been the specific concern of experimentation by Lumley (7, 8), Ruch (9), Spragg (11), and Blodgett and McCutchan (1, 2). Blodgett and McCutchan (1) have suggested that "anticipation in the semi-linear elevated maze results from failure of the rat to discriminate between the place of the final choice-point and that of the immediately preceding choice-point where anticipation occurs" (p. 351). They conclude that "anticipatory behavior is an aspect of the absence of place learning."

B. MATERIALS AND METHODS

1. *Maze and Room*

The floor plan of the maze and room are shown in Figure 1. The individual units of the maze were painted flat black and were of the following dimensions: 24 inches by 2 inches, and were 15 inches in height. The individual units were shifted frequently in the pattern; the specific location and direction of the maze, however, remained constant.

The experimental room was 20 feet by 20 feet by 10 feet and was light-proof. The circular dome in the experimental room was designed to give a homogeneous appearance. The frame, constructed of white pine, was covered with a fine steel screen mesh.

The source of illumination for the present experiment was one $7\frac{1}{2}$ -watt blue-colored light bulb, suspended 12 inches from the dome-ceiling. The light rays were deflected upward by an inverted opaque conical deflector. The upward deflected rays were then further reduced in intensity by placing a double layer of dark-colored medium weave cloth at the opening of the deflector.

2. *Animals*

The animals were descendants of Wistar stock bred in the Comparative Psychology Laboratory of the University of Texas. They were between 90 and 120 days old at the beginning of the study and were experimentally

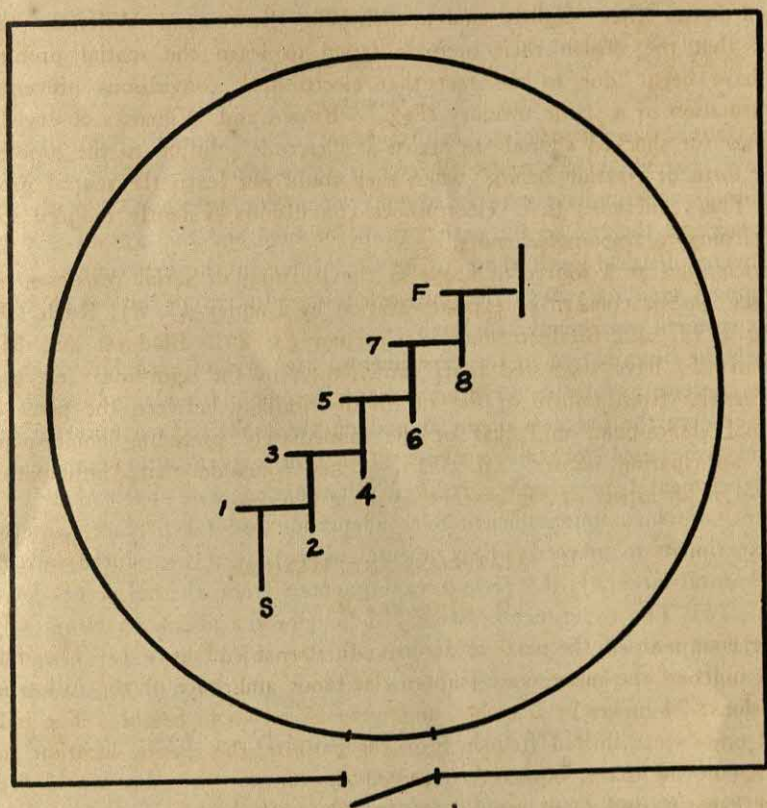


FIGURE 1
MAZE AND EXPERIMENTAL SETTING

naïve. The daily rations consisted of a mash which was mixed with water before feeding.² Each animal received its entire daily ration in an individual food box at the end of the maze. No other feedings were made, although water was available in the group cages at all times.

3. Scoring

Body length entry into a cul was recorded an error. Errors made during retracing were not recorded as errors. Retrace scores and running time

²The formula for this mash consisted of the following ingredients (by weight): 5 parts yellow cornmeal, 1 part linseed meal, 1 part powdered milk, 1 part ground alfalfa. For each 5 pounds of cornmeal, about 50 grams of salt and 50 grams of calcium carbonate were added. Every other day 10 drops of codliver oil U. S. P. were added to each ration.

records were kept. Animals were not permitted to retrace for more than one unit on any attempted retrace.

4. Procedure

To familiarize the subjects with general maze-running procedure, preliminary training on a straight runway was provided for five days. Three individual maze units were placed end-to-end and a goal box with food was placed at the end of the path. Both the food and the type of goal box used were identical with those to be used later in the experiment. This preliminary training was carried out in a room adjacent to the one in which the experiment was conducted later.

Since the illumination in the experimental area was of such low intensity, it was necessary for one experimenter to become dark-adapted before he could observe the progress of an animal on the maze. Two different procedures were used for this purpose. (a) One experimenter remained in the experimental area under reduced illumination for approximately 15 minutes, at which time sufficient dark adaptation had taken place to permit the experiment to proceed. Experimental animals were introduced into the experimental area by the second experimenter from the adjacent colony room. (b) The experimenter wore red goggles for about 15 minutes before the experiment began. This procedure enabled the experimenter to go in and out of the experimental area without additional dark adaptation.

5. Group I (Control Group — $N = 16$)

The animals in the Control Group were given one run a day for 15 days. Cul errors, retrace errors and time records were kept. The animals in Group I were not given electroconvulsive shocks.

6. Group II (Daily Shock Group — $N = 16$)

The animals in this group were trained in the same manner as those in Group I, with one exception: following each day's run, and subsequent to its eating period, each animal was administered an electroconvulsive shock of 50 ma. intensity for 0.20 seconds (see 6, p. 40-42). The animals were then returned to their home cages.

7. Group III (Twenty Days Shock Before Training — $N = 10$)

The animals in Group III were given one electroconvulsive shock daily for 20 days before training began on the maze. They were not shocked during the training period, their daily routine being the same as that of the Control Group. The current used for shock was the same as that used for Group II.

C. RESULTS

1. Error Scores

The ipso-final (Culs 1, 3, 5, and 7) error scores are presented graphically in Figure 2 for all three groups in terms of percentages. An examination of these data will show a difference in the performance of the three groups. The error score curve for Group I (non-shock) declines gradually

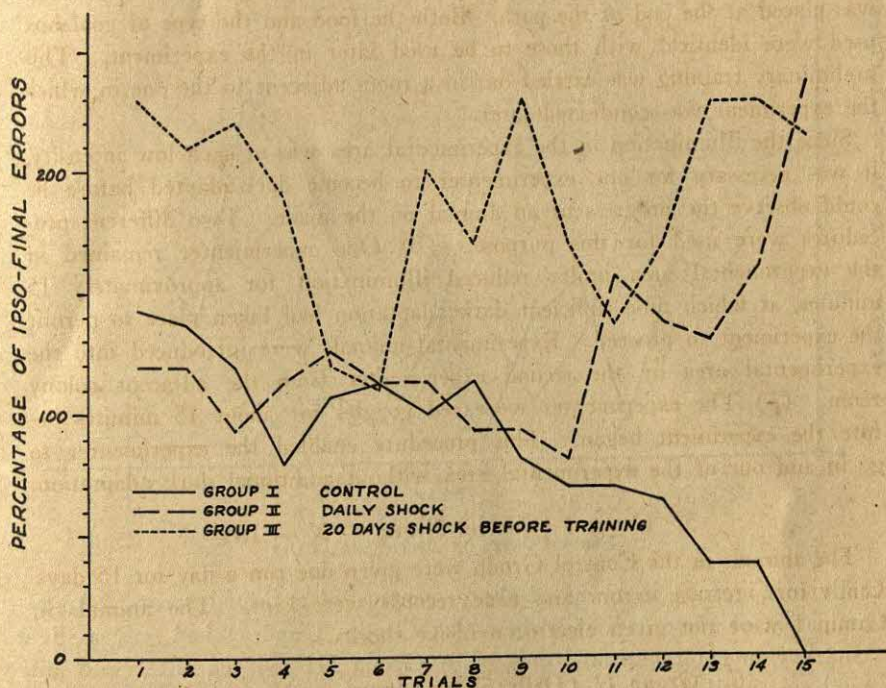


FIGURE 2
ERROR CURVES ON IPSO-FINAL CULS

and reaches zero on the fifteenth trial. The error score curve for Group II (daily shock) is very similar to that of Group I until the tenth trial, after which time the two curves separate, with errors increasing for Group II and decreasing for Group I. Ipso-final errors are much greater for Group II on the fifteenth day than on the first day of training. An examination of Figure 2 will show that the ipso-final error curve for Group III (20 days shock before training) may be characterized best as an erratic pattern.

The anti-final (Culs 2, 4, 6, and 8) error scores are presented graphi-

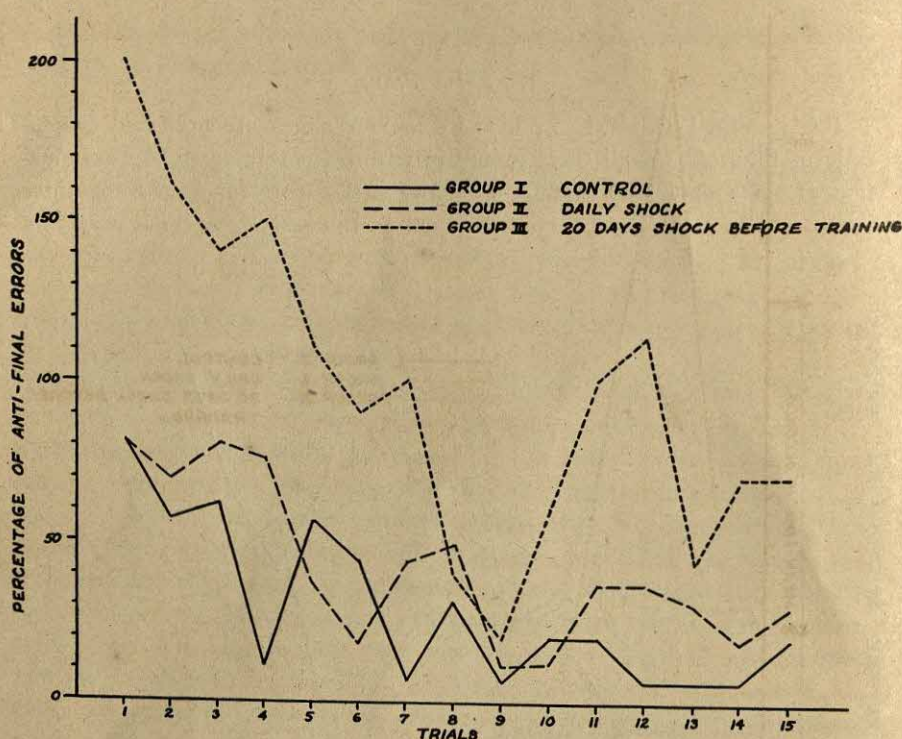


FIGURE 3
ERROR CURVES ON ANTI-FINAL CULS

cally in Figure 3 in terms of the percentage of anti-final errors. An examination of these data will show that Groups I and II had only a slight tendency to enter anti-final culs while Group III reduced entry into anti-final culs until the ninth trial, after which there was a sharp rise and then another decline.

2. Time Scores

Pronounced differences in the amount of time required to run the maze were observed between Groups I and II, and Group III. These differences were especially marked at the beginning of the experiment, but were definitely present throughout as may be clearly seen in Figure 4.

3. Retrace Scores

Retrace curves are presented for all three groups in Figure 5. These data show that Group II had a somewhat stronger tendency to retrace than

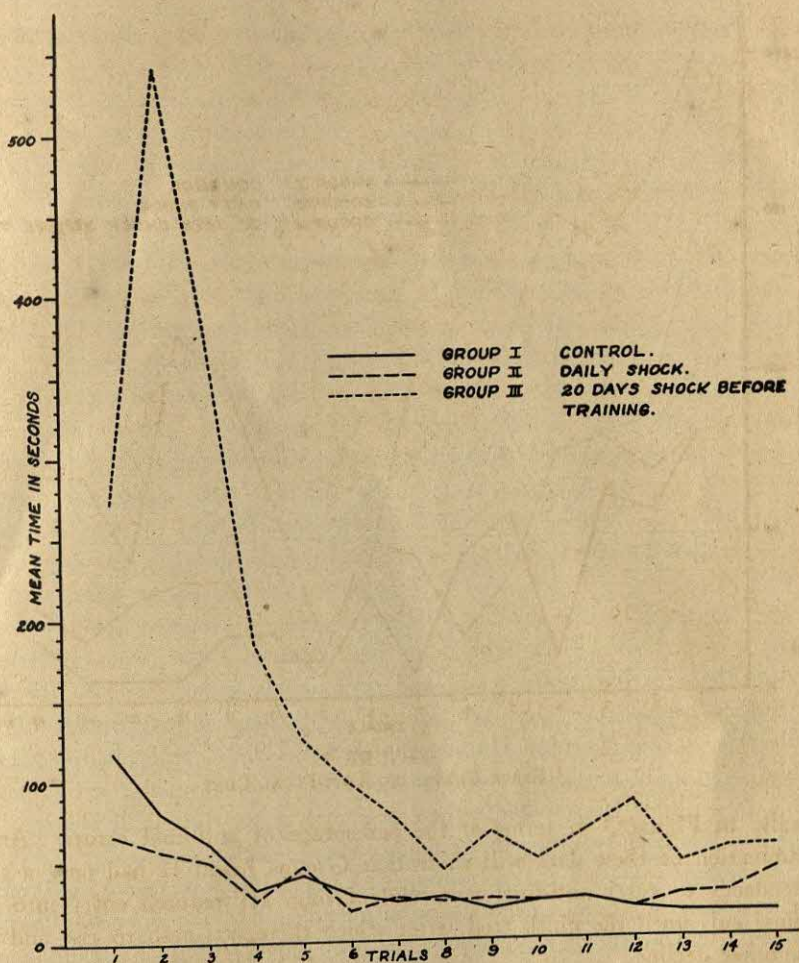


FIGURE 4
TIME CURVES

did Group I, but neither group had as strong a tendency to retrace as did Group III. After an initial high percentage of retrace errors for Group III a rapid drop in the error percentage was effected.

4. Observations

The non-shock group (Group I) behaved in the manner characteristic of normal rats on a maze problem. They were docile, easily managed,

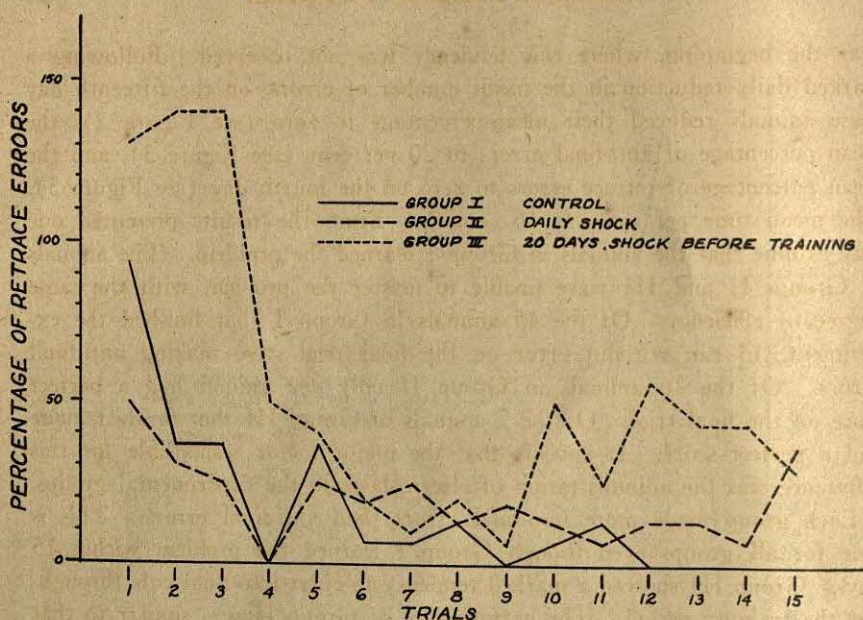


FIGURE 5
RETTRACE ERRORS

and as the experiment progressed and the animals became used to the routine, they became even less reluctant to being handled. Members of the daily shock group (Group II) present an entirely different picture. They were abnormally sensitive to sudden noises or sudden rapid movements by other rats. This was especially the case shortly after their daily shock when they had just been returned to their home cages.

In the early days of maze training of Group III, it was not uncommon for an animal to fail to negotiate a turn and fall to the floor. This phenomenon was prevalent especially at the beginning of the maze training and limited only to the animals that made up Group III. When Group III animals were placed at the starting point on the maze, it was not uncommon for them to urinate and/or defecate. This behavior was also observed at other points along the maze whenever the animal had unduly lengthened the distance and the time from the starting point to the food box by entering an excessive number of culs and had made several retraces.

D. DISCUSSION OF RESULTS

The animals in Group I (non-shock) tended to decrease the mean percentage of errors from the beginning. There was only one brief interval,

near the beginning, where this tendency was not observed. Following a marked daily reduction in the mean number of errors, on the fifteenth day these animals reduced their mean-percentage to zero (see Figure 2), the mean percentage of anti-final errors to 20 per cent (see Figure 3), and the mean percentage of retrace errors to zero on the fourth day (see Figure 5). The mean time per rat was 16 seconds. From the results presented one may assume that the animals in Group I learned the problem. The animals in Groups II and III were unable to master the problem with the same degree of efficiency. Of the 15 animals in Group I that finished the experiment, 13 ran without error on the final trial, two making anti-final errors. Of the 16 animals in Group II only *one* animal had a perfect score on the final trial. Of the 7 animals of Group III that finished, none had a perfect score. It appears that the major factor responsible for this difference was the administration of electroshock to the experimental groups.

Each group made more ipso-final errors than anti-final errors. This is true for all groups even though Group I learned the problem within 15 days. Group III showed a marked tendency to enter ipso-final culs throughout the training period. The performance of Group II was similar to that of Group I for the first 10 days, after which there was a sharp increase in ipso-final errors for Group II and a decrease in ipso-final errors for Group I. Apparently the electroshock animals were unable to *inhibit* the making of anticipatory responses in the semi-linear maze. Brown and Wilbanks (4) have observed that electroshock convulsions tend to reduce "spatial learning to response learning." The present problem tended to accent the difficulty of the ipso-final culs over the anti-final culs. The animals in Group I (non-shock) were able to inhibit the making of ipso-final errors, but the animals in the two shock groups were never able to inhibit entry into ipso-final culs. The animals in each group developed very early a tendency to enter ipso-final culs (to make anticipatory responses). The shock animals were never able to inhibit this tendency. Cul entry by the shock animals was usually to the end of the cul and not just body length or partial entrance. As learning developed on the part of Group I numerous VTE's were observed at the bifurcations. Such behavior was not observed in Groups II and III. The animals in Groups II and III learned that a left turn was essential for mastery of the problem but never learned when not to turn left. Group I was able to inhibit a left turn in sections identical in pattern to the final section but Groups II and III were not. Since Groups II and III were unable to inhibit a left turning tendency, their errors increased during the training period in ipso-final culs. The

animals in Groups II and III learned a simple left turning response which dominated their entire performance. They were never able to give the response only at a specific bifurcation in the maze.

E. CONCLUSIONS

The foregoing study has demonstrated that:

1. The non-shock animals were able to learn a semi-linear maze in very dim illumination with homogeneous surroundings. Their errors in ipsosinal culs were much greater than in anti-final culs before learning developed.
2. The daily electroshock animals (Group II) were unable to inhibit the making of anticipatory responses in a semi-linear maze with dim illumination with homogeneous surroundings.
3. The behavior of the 20-day preshock group (Group III) was very erratic. They were also unable to inhibit the making of anticipatory responses in the semi-linear maze used.
4. The two shock groups did not learn the semi-linear maze during the 15-day training period.

REFERENCES

1. BLODGETT, H. C., & MCCUTCHAN, K. Place learning and anticipation: I. Anticipation as failure of place discrimination. *J. Comp. Psychol.*, 1945, **39**, 351-357.
2. ———. Anticipation and place learning: II. Anticipation in a visually homogeneous environment. *J. Comp. Psychol.*, 1947, **40**, 16-175.
3. BROWN, W. L., & FOSMIRE, F. R. The effect of electroconvulsive shock on the learning of a spatial maze. *J. Genet. Psychol.*, (In press.)
4. BROWN, W. L., & WILBANKS, W. A. The effect of different periods of electroconvulsive shock on spatial learning. *J. Genet. Psychol.*, (In press.)
5. FOSMIRE, F. R., & BROWN, W. L. A preliminary study of the effect of electroconvulsive shock on spatial learning. *J. Genet. Psychol.*, (In press.)
6. HAYES, K. J. Cognitive and emotional effects of electroconvulsive shock in rats. *J. Comp. & Physiol. Psychol.*, 1948, **41**, 40-61.
7. LUMLEY, F. H. Anticipation as a factor in serial and maze learning. *J. Exper. Psychol.*, 1932, **15**, 331-342.
8. ———. Anticipation of correct responses as a source of error in the learning of serial responses. *J. Exper. Psychol.*, 1932, **15**, 195-205.
9. RUCH, F. L. Experimental studies of the factors influencing the difficulty of blind alleys in linear mazes: II. Generalized turning habits; III: Is there an anticipatory tendency in maze learning? *J. Comp. Psychol.*, 1935, **20**, 35-52 and 113-124.
10. RUSSELL, R. W. Effects of electroshock convulsion on learning and retention in rats as a function of difficulty of the task. *J. Comp. & Physiol. Psychol.*, 1949, **42**, 137-142.
11. SPRAGG, S. D. S. Anticipation in maze errors. *J. Comp. Psychol.*, 1933, **15**, 313-329.

Department of Psychology
University of Texas
Austin 12, Texas

A REVIEW OF THE LITERATURE ON THE EFFECTS OF BILINGUALISM UPON THE MEASUREMENT OF INTELLIGENCE*¹

Department of Education, Fordham University

NATALIE T. DARCY

A. THE PROBLEM

A comprehensive survey of the literature on the intellectual and educational effects of bilingualism would include many early studies² which, while interesting from an historical point of view, were completely lacking in experimental method. Such studies, therefore, would have little significance for those interested in further experimental work in this field. Accordingly, this article will be limited to a review of a representative number of studies which, with an effort toward controlled experimentation, dealt with the effects of bilingualism upon the measurement of intelligence. The selection was made with the purpose of assisting the future investigator and from this point of view the investigation will be classified under the following headings:

1. Studies in which bilingualism was found to have a favorable effect upon the measurement of intelligence.
2. Studies in which bilingualism was found to have an unfavorable effect upon the measurement of intelligence.
3. Studies in which bilingualism was found to have no effect upon the measurement of intelligence.

The classification is one of convenience and cannot, because of the nature of the studies and the nature of the subject, avoid some overlapping. The reviews, however, have been presented in some detail and the discerning reader will have little difficulty in making any necessary qualifications which may meet his particular needs.

*Received in the Editorial Office on February 27, 1951.

¹This review is based upon the review of the literature in the author's dissertation completed under the direction of Professor William A. Kelly and accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Fordham University, June, 1945.

²The reader who wishes a more comprehensive knowledge of the studies in this field will probably find the bibliography at the close of this article adequate for his needs.



B. STUDIES IN WHICH BILINGUALISM WAS FOUND TO HAVE A FAVORABLE EFFECT UPON THE MEASUREMENT OF INTELLIGENCE

Davies and Hughes (28) conducted an investigation to determine what differences, if any, were observed between Jewish and non-Jewish children in intelligence, English, and arithmetic. All the pupils from 8 to 14 years of age, in three carefully selected London schools, were given the Northumberland Standardized Tests in Intelligence, English, and Arithmetic. The total number of subjects who received these tests was 1894. The only information which was given with respect to the control of variables was expressed in the statement that "the three schools in which the study was made, were chosen in an attempt to rule out differences due to variations in school teaching and home training, except in so far as these latter are due to racial customs and traditions" (28, p. 136). The degree of bilingualism, which might be presumed to have been included in "the racial customs and traditions" of at least part of the Jewish group, was not noted.

The authors concluded that the superiority of the Jewish children over the non-Jewish children in the two groups which were studied was definitely marked. The Jewish group was found to be one year ahead of the non-Jewish group in intelligence and arithmetic and one and one-half years ahead of the non-Jewish group in English. The Jewish girls were one year ahead of the non-Jewish girls in English and slightly less than a year ahead of the non-Jewish girls in intelligence and arithmetic.

Stark (101) measured the general intelligence of 271 bilinguals in five "all-Irish schools" in the city of Dublin and 271 monoglots in four "non-all-Irish schools" in the same city. The subjects, who ranged in age from 10 to 12 years, were given the Dawson Mental Tests, which are group intelligence tests, and the Passalong Test, which is an individual performance test.

An Irish translation of the Dawson Mental Test, Form *B*, was made and it was found to have a correlation of $.84 \pm .02$ with the English form of the test. The bilinguals were given the Irish form while the monoglots received the English form. A difference in mean score of five points, in favor of the monoglots, was found at ages 11 and 12. The differences in mean score between the two language groups at earlier ages were found to be negligible.

When the Dawson Mental Test, Form *A*, was administered in English to 104 monoglots and to 65 bilinguals, who had been selected at random from the original language groups who had received the Dawson Mental

Test, Form *B*, a comparison of mean scores showed the bilinguals at age 10 to be superior to the monoglots of the same age by 13 points. At age 11, the mean score of the bilinguals was seven points higher than the mean score of the monoglots. Both of these differences were found to be statistically significant.

The Passalong Test was administered to 41 monoglots and 41 bilinguals and no significant difference was found to exist between the two language groups.

As a result of these findings, Stark concluded that children of "innate verbal facility" may find early bilingualism an asset to their mental development. It must be remembered, when examining Stark's results, however, that aside from selecting subjects of the same age range who attended primary schools in one city, the variables such as socio-economic status, sex, and degrees of bilingualism, which might have affected the results, were not adequately controlled.

C. STUDIES IN WHICH BILINGUALISM WAS FOUND TO HAVE AN UNFAVORABLE EFFECT UPON THE MEASUREMENT OF INTELLIGENCE

Saer (82-84) conducted a series of experimental investigations to determine the effect of bilingualism on the measurement of intelligence. The subjects included 1,400 children from five rural and two urban districts in Wales. In six of these districts Welsh was the mother tongue but the children of the rural districts learned English at school while those of the urban districts learned English at school and English was also the language of their play activities. The seventh was a rural district in which English was the mother tongue. Information was obtained concerning the socio-economic backgrounds, home language, and age of each child but no attempt was made to match monoglots with bilinguals with respect to these factors.

The 1916 Stanford-Binet Scale and those tests in the 1911 Binet Scale and Burt's Revision which had been omitted from the Stanford Revision, were administered to all the subjects. These tests were translated into the Welsh language for those who used Welsh as their mother tongue since the author felt that the mother tongue is the best oral medium by which a just estimate of a child's mental capacity can be gained. The results on these tests were as follows:

Urban bilingual group—median <i>IQ</i>	= 100
Urban monoglot group—median <i>IQ</i>	= 99
Rural bilingual group—median <i>IQ</i>	= 86
Rural monoglot group—median <i>IQ</i>	= 96

Saer presented tables listing the median mental ages and the frequencies of *IQ*'s. Both of these tables indicated the significant inferiority of the rural bilinguists and suggested that this inferiority became consistently greater in degree with each year, from 7 to 11 years of age.

The subjects were given also a vocabulary test consisting of a 100-word list from Anwyll's Welsh Dictionary. In the rural districts, a sharp rise was found at the age of 10 years in the range of the children's vocabulary in both English and Welsh. In the case of monoglot English-speaking children in rural districts, the most considerable rise in vocabulary occurred at eight years of age while the greatest rise in the bilingual group occurred at the 10-year age level. In the urban districts, the sharpest rise occurred between 9 and 10 years in the case of both monoglots and bilinguists. The mean range of vocabulary of monolingual children was higher than that of the bilinguists in both English and Welsh. The author accounted for this by stating that mental confusion occurs in the bilingual children more often than in the monoglots.

In a composition test in which the children were asked to write of a dream and also about anything they liked best, it was found that the bilinguists showed greater clarity and correctness when writing in Welsh than when writing in English.

The results of a test of dextrality, in which 679 bilinguists and 281 monoglots from 8 to 11 years of age, were asked to show their right hand, left ear, or in a picture to show the right hand, ear, and foot, the bilinguists in both rural and urban districts were found to be much more confused than the monoglots. The author suggested that the confusion had been carried over from the brain area concerned with language to the related specialized area connected with the use of the right hand in gesture or in writing.

A rhythm test was given to 339 unselected subjects in an urban school. They were given a tapping exercise in which the monoglots were found to be superior to the bilinguists at 8, 11, and 12 years of age but the bilinguists were superior to the monoglots at 7, 9, and 10 years of age. However, when the rhythms were sung to "la," the monoglots were found to be superior at all ages.

A verbal intelligence test was administered to 939 students in the University College of Wales. The monolingual students showed a "considerable superiority" over the bilingual students from the rural districts of Wales. The difference in the measurement of intelligence between the monolingual and bilingual students from urban and industrial districts was found to be

"inconsiderable," however, and Saer stated that the difference seemed to be of a permanent nature since it was seen to persist in students throughout their university careers.

As a result of his experiments, Saer concluded that children who become bilingual at an early age, by learning the second language during their play and in contact with other children, have an advantage over those who learn the second language at school. The fact that the Stanford-Binet Scale was translated into the Welsh language for the children who spoke Welsh at home must be taken into account in viewing Saer's results for the translation of a standardized test is not an equivalent of the test in the language in which it was originally standardized.

Pintner and Keller (72) reported an investigation in which the children in the kindergarten, first, and second grades of three schools located in Youngstown, Ohio, were divided into an English-speaking group and a foreign group. This language classification was made on the basis of the parents' nationalities. The English-speaking group included the American-white, Negro, English, Canadian, Scotch, Irish, and Welsh populations while the foreign group consisted mainly of children of Italian and Spanish parentage. There were 367 children in the English-speaking group and 674 children in the foreign group.

All subjects were given a special revision of the 1916 Stanford-Binet Intelligence Scale. This special revision was prepared and administered by the Service Bureau of Youngstown and was found to have a correlation of .97 with the 1916 Stanford-Binet Intelligence Scale. The average *IQ* of the English-speaking group on this revised Binet Scale was 92 while the average *IQ* of the foreign group was 84.

The Pintner Non-Language Group Test was administered to the second grade children in one of these schools and on this test, the English-speaking group attained an average *IQ* of 109 whereas, the foreign group received an average *IQ* of 103. It was also found that 73 per cent of the children in the English-speaking group and 82 per cent of the children in the foreign group made higher *IQ*'s on the Pintner Non-Language Test than on the Binet Revision.

A comparison was then made between the results obtained by a group of children who were given the 1916 Revision of the Stanford-Binet Scale and a series of performance tests including at least three of the following: Healy Construction Puzzle, the Wittmer Cylinder Test, the Pintner Cube Test, and the Mare and Foal Test. The correlations between the Binet and the performance tests for this group were .64 for the English-speaking group and .48 for the foreign group.

The authors concluded that children from homes in which a foreign language is spoken, receive lower scores on revisions of the Stanford-Binet Scale than they do in tests in which a minimum of English is used and that these children may suffer a serious handicap if classified with English-speaking children on the basis of *MA*'s achieved only on revisions of the Binet Scale.

Brown (16) studied the scores which had been made on the Stanford-Binet Scale by 913 children of nine different nationalities, all of whose parents were foreign-born. He found that bilingualism presented the chief difficulty in the testing of these children although, in general, pupils who had attended an American school for one or two years received no lower scores on the Stanford-Binet Scale than they received when this scale had been translated into their native tongues. In some cases, however, children who spoke the English language fairly well, had mental ages from 6 to 18 months higher when tested by the Stanford-Binet Scale in their native tongues than when tested by the Stanford-Binet Scale in English. Caution should be exercised in accepting these results, however, since the translation of the Stanford-Binet Scale is not an equivalent of the scale in English.

Colvin and Allen (22) reported an investigation of 50 children of native born parentage and 50 children of Italian parentage who were attending grades five through eight in the public schools of Providence, Rhode Island. The ages of the American group ranged from 11 to 15 years. All the subjects were given the National Intelligence Test and the 1916 Stanford-Binet Scale. The average *IQ* on the Stanford-Binet Scale for the American group was 92 while the Italian group received an average *IQ* of 91 on this scale. On the National Intelligence Test, the Americans received an average *IQ* of 85 and the Italians an average of 76.

These results showed that the children in both groups had a tendency to be rated lower in *IQ* by the National Intelligence Test than by the 1916 Stanford-Binet Scale and this difference was considerably greater for the Italian group. The authors concluded that an individual test constitutes a more accurate measure of intelligence than a group test and that the linguistic factor seems to be less important in the Stanford-Binet Scale than in the National Intelligence Test. With respect to the weight which is to be given to the linguistic factor when viewing intelligence test results they state that "while verbal ability may raise intelligence scores in some instances above the level of the actual intelligence of the person examined, its most marked effect is noted under the condition when the lack of such a facility unduly lowers such scores" (22, p. 3).

Pintner (69), in a study which sought to determine whether or not a

verbal group test is a valid measure of the intelligence of foreign children, administered the National Intelligence Test, Scale *A*, Form *I*, and the Pintner Non-Language Test to all the children in the third and fourth grades of a New York City school. The children were divided into foreign and American groups. The foreign group consisted of Italians, Germans, and Poles but the number of German and Polish subjects was so small that it was of little consequence.

The median *MA* for the total foreign group on the Pintner Non-Language Test was nine years four months as against a median *MA* of eight years nine months on the National Intelligence Test, while the median *MA* for the American group on the Pintner Non-Language Test was nine years four months as against a median *MA* of nine years no months on the National Intelligence Test.

In comparing the separate nationality groups, Pintner found that the median *MA*'s on the Pintner Non-Language Test for the Germans and Poles were above the median *MA* for the Americans; whereas, the median *MA* for the Italians on the non-language test was below that of the Americans. For all the foreign children combined, the median *MA* on the Pintner Non-Language Test was the same as the median *MA* of the Americans on that test.

In studying the percentage of foreign children who reached or exceeded the median *MA* of the American group, Pintner found no difference between the foreign group as a whole and the American group on the Pintner Non-Language Test. On the National Intelligence Test, however, only 37 per cent of the foreign group reached the median *MA* of the American group. When comparing the Italian group with the American group, it was found that 43 per cent of the Italians reached or exceeded the median *MA* of the American group on the Pintner Non-Language Test whereas, on the National Intelligence Test only 36 per cent of the Italians reached or exceeded the median *MA* of the Americans.

The author concluded that caution should be exercised in drawing conclusions concerning the intelligence of foreign children when verbal tests of intelligence are used as the sole criterion of measurement.

Smith (92), in an investigation in South Wales, studied bilingual and monolingual subjects from 8 to 11 years of age, who were attending four rural schools. The subjects were given tests in free composition, a mutilated passage, word-building and analogies, three times within a period of two years. In comparing the coefficients of correlation between successive tests, the investigator found that the free composition and mutilated passage

tests were more reliable than the word-building and analogies, three times within a period of two years. In comparing the coefficients of correlation between successive tests, the investigator found that the free composition and mutilated passage tests were more reliable than the word-building and analogies tests.

When the results in the four schools were summarized and averaged, the monoglots were found to be superior to the bilinguals on all four tests. The author concluded that bilingualism seemed to be an intellectual disadvantage under the organization of schools in South Wales at the time of the study. The validity of Smith's results might be questioned since the two language groups were not paired with respect to socio-economic status or age and also because of the lack of objective scores for the composition test.

Graham's (42) study was based on all the objective data from intelligence tests that were available in the files of the Habit clinics of the Massachusetts Division of Mental Hygiene in the Fall of 1924. The records of 60 Americans, 43 Italians, and 47 Jews were studied. Since most of the subjects were American-born the racial characteristics were determined by parentage only. The tests used were the Stanford-Binet Scale and a series of performance tests including Porteus Maze, Knox Cube, Healy Pictorial Completion I, Mare and Foal, Diagonal Form Board, Triangle Form Board, Healy Construction A, and the Manikin Test. Not all of the subjects were given all of the tests.

The results of the Stanford-Binet were as follows:

Americans—Mean <i>IQ</i>	99 σ 24
Jews—Mean <i>IQ</i>	105 σ 23
Italians—Mean <i>IQ</i>	85 σ 16

The author stated that if there were no language handicaps present, these results would have given a fair estimate of the general intelligence of the three groups.

In correlating the scores made on the performance tests with those made on the Stanford-Binet, it was found that none of the performance tests on which the Italian subjects had been tested, seemed to have measured the same type of intelligence as the Stanford-Binet Scale had measured. Only the Cube Test correlated highly with the Stanford-Binet Scale as far as the Jewish subjects were concerned, while all correlations between performance tests and the Stanford-Binet Scale were fairly high for the Americans. Graham explained the lack of correlation between performance and verbal tests for the foreign groups on the basis of language handicap since

the performance test results were found to be positively correlated with the chronological age of Italian groups when the Stanford-Binet results were held constant. Furthermore, the racial relationship proved to be the same when the performance tests were used as the measures of mental ability.

The results of this investigation may be questioned, however, since the author made no mention of having controlled such variables as age or socio-economic status when comparing the nationality groups.

Wang (105), in a study which was conducted at Ohio State University, paired native-born white students with 34 Chinese students, 158 Negroid students, and 45 Russian students with respect to college attended, class, sex, and age. All subjects were given the Ohio State University Intelligence Test. The results showed that the native-born subjects averaged nine steps above the Chinese in the General Information Test. On the Number Series of the scale, however, the Chinese averaged 8.5 steps above the native-born subjects. There were no outstanding differences between the scores of the Russians and native-born white students or between the scores of the Negroes and Native-born white students. Wang stated that the Chinese were handicapped by the language difficulty and that in the case of the Negroes and the Russians, the language factor did not seem to be an appreciable one. While admitting that his subjects were selected and not representative of these racial groups in the country at large, he concluded, as a result of his investigation, that "one should not draw any definite conclusion concerning the relative intelligence of racial groups not thoroughly Americanized and Americans, unless there is some control of the language factor. Non-verbal or performance tests might be more equally valid for several or all races" (105, p. 106).

Mead (65) administered the Otis Group Intelligence Test, Forms *A* and *B*, to 276 native-born children of Italian parentage and to native-born children of American parentage, who were in grades 6 to 10 in several New Jersey public schools. The results were analyzed in the light of available norms on the Otis Group Intelligence Test and some biographical data for the Italian subjects which consisted of the language or languages spoken in the home, the dates of immigration of the parents, and the social status of the parents. Since a correlation coefficient of .82 was found between Forms *A* and *B* of the Otis Group Intelligence Test, the results of Form *A* only were used in the analysis.

The difference between the mean indices of brightness on Form *A* of the Otis Group Intelligence Test for the American and Italian groups was $27.33 \pm .21$ in favor of the Americans. Only 6 per cent of the Italians

exceeded the median index of brightness of the Americans and only 7 per cent of the Italians exceeded the median *IQ* of the Americans. The difference between the mean indices of brightness for the Italian boys and girls was $6.9 \pm .2$ in favor of the girls, and the difference between the mean indices of brightness for the American boys and girls was $8.5 \pm .36$ in favor of the girls.

When the Italians were classified into four social-status groups, it was found that only 12 children or .05 per cent of the group could be classed as fully comparable with the average American child in the community. The author states that this factor, as well as the language factor and the length of time the fathers had spent in the country, influenced the scores of the Italian children. He concluded, therefore, that classification of foreign children in schools where they have to compete with American children on the basis of group intelligence tests results alone, is not a just evaluation of the foreign child's innate capacity. Since this investigation did not control those factors, other than the language factor, which may have had some influence on the scores of the Italian children, it is impossible to conclude just how much weight the language factor had in lowering the scores of the foreign group.

Rigg (78) investigated the results of one of the St. Louis testing programs in an attempt to determine the effect of a language handicap upon the results of intelligence and achievement tests. The children who were tested were in grades three through eight of the public schools of St. Louis. The "foreign" children were identified by their answers to the question, "What foreign language is spoken in your home?" The tests which had been administered are the National Intelligence Test, Scale *A*, Form *I*; the Thorndike-McCall Reading Test, Form *I* and the Woody-McCall Arithmetic Fundamental Test, Form *I*. For the purposes of the present investigation, however, this review will be confined to a report of the differences in the results on the National Intelligence Test between the 8,130 native-born white children and the 140 Italian children. On this test the median *IQ* of the native-whites was 104.85 with a *PE* of .17 whereas, the median *IQ* of the Italian group was 91.43 with a *PE* of 1.13. This indicated a significant difference in favor of the native-born whites on this intelligence test.

Rigg admitted that he did not take the social status of the groups into account nor were other variables controlled to any great extent in this study. In view of the data, however, the author concluded that the language handicap may explain the lower standing of most of the foreign groups on the National Intelligence Test but that it is insufficient to explain the large variation in the case of the Italians.

Jamieson and Sandiford (53) studied approximately 700 Indian subjects in Southern Ontario. All of the subjects were able to speak the English language, but the authors stated that the subjects were not able to use English with the facility of native-born white children of corresponding social status. It was also found that 45 per cent of the subjects could speak an Indian dialect and many of the remaining 55 per cent could understand, although they could not speak the dialect. When measured by the Chapman Socio-economic Scale, the subjects received an average score of 13 against a norm of 56.

The National Intelligence Test, Scale *A*, Form *I*, the Pintner-Cunningham Primary Mental Test, the Pintner Non-Language Scale and the Pintner-Patterson Performance Scale were administered. The results were as follows:

National Intelligence Test, Scale <i>A</i> , Form <i>I</i> —median <i>IQ</i>	= 80
Pintner-Cunningham Primary Mental Test—median <i>IQ</i>	= 78
Pintner Non-Language Scale—median <i>IQ</i>	= 97
Pintner-Patterson Performance Scale—median <i>IQ</i>	= 92

The authors stated that these results pointed to a severe language handicap on the verbal tests since the monoglots surpassed the bilinguals on all tests except the Pintner-Patterson Scale of Performance Tests, and the difference between the two groups in this test was not so significant as the differences between the two groups on the other three tests.

Yoshioka (110) administered the National Intelligence Test, Scale *A*, Form *I*, in English and in Japanese to 38 Japanese children in California. The subjects, who ranged from 9 to 15 years in age, were divided into two groups. Group *A* consisted of 17 children in grades three through five and Group *B* was composed of 21 children in grades six through eight. Group *A* was given the Japanese form of the test on the first day of the testing and the English form on the second day; the order was reversed for Group *B*.

In both forms of the intelligence test, the children who were the subjects of this investigation scored far below the American and Japanese norms. Yoshioka explained this on the basis of the language difficulty caused by bilingualism and stated that the low correlation between the Japanese and English forms gave further proof of the language difficulty, although the investigator added that the children on whom the norms had been based for the Japanese form of the test were of higher socio-economic status than the subjects of his investigation. The correlation between the two forms of the test was higher for Group *B* than the correlation between the two forms for

Group *A*. Yoshioka attributed this to the fact that the older subjects were better bilinguals, and added: "It is suggested that bilingualism in young children is a hardship and devoid of apparent advantage, because bilingualism appears to require a certain degree of mental maturation for its successful mastery" (110, p. 476). The small number of cases in this study minimized the validity of its conclusions.

Pintner (70) conducted an investigation in which the Pintner-Cunningham Primary Mental Test and the Pintner Primary Non-Language Test were administered to 430 children in Grades 1A and 1B in three elementary schools in New York City. The surnames of the subjects, or the judgment as to language background by their teachers or Pintner's assistants in the investigation, determined whether or not those who were tested should be placed in the bilingual or monolingual groups.

In School *A*, where the second language of the bilingual group was Italian, it was found that on the Pintner-Cunningham Test, the mean score of the bilingual group was 8.3 points less than the mean score of the monolingual group. In terms of standard deviation, the bilingual group was .89 below the monolingual group. On the Pintner Non-Language Test, however, the bilingual group had a mean score only 4.02 points lower than the mean score of the monolingual group and in terms of standard deviation, the bilingual children on this test were .31 below the monoglots.

In School *B*, where the second language was Bohemian, the bilingual group exceeded the monolingual group on both tests.

In School *C* no difference was found, but the author stated that this was probably due to the fact that no real division existed in the language background of the pupils in this school.

Pintner concluded from these results that caution should be exercised when comparing bilingual and monolingual children on the basis of verbal intelligence tests. It must be noted, however, that neither the nationality groups nor the language groups were matched as to socio-economic status and there was only a very general and subjective determination of bilingualism.

Sanchez (85) tested 45 Spanish speaking children in grades three through eight in the public schools of New Mexico. The Stanford Achievement Test, Forms *A*, *B*, and *V*, and the Haggerty Intelligence Tests, Delta 1 and Delta 2, were administered to the subjects at four intervals between December, 1928, and April, 1930. The average mental age of the group increased with each successive test (Table 1).

It will be seen that the greatest increase in average *MA* occurred between the first and second testings while the period between the second and

TABLE 1

Data of test	Average <i>MA</i> in months
December, 1928	101.0
April, 1929	118.0
November, 1929	128.9
April, 1930	143.0

third testings showed the least increase. This period included three months of summer vacation. No general tendency toward an increase of *IQ* with an increase in age was found. The author attributed the gains which were made in average mental age to the increased ability in reading and the added language facility which had been achieved during the intervals between the two tests.

Barke's (6) investigation had for its objectives the comparison of the general intelligence of pupils in certain bilingual and monolingual schools in South Wales. The author applied the term bilingual to "schools in which the mother tongue is used almost exclusively with the younger children but a second language becomes progressively the chief medium of instruction in the upper classes" (6, p. 237).

The 395 bilingual subjects for the investigation ranged from 10 to 14 years of age and were selected from three mixed bilingual schools near the mining district of South Wales where Welsh is the dominant language. In answer to a questionnaire, 86.2 per cent of the subjects replied that they spoke Welsh at home.

The monolingual group was composed of 302 subjects from 10 to 14 years of age, who were chosen from the children in two schools in a mining district of the same country. In this district English was the dominant language in school and at home but there was a compulsory 45-minute Welsh lesson in each of these schools every day. In answer to a questionnaire, only 3.4 per cent of the boys and 3.9 per cent of the girls replied that they spoke Welsh at home. The home environment of the bilinguals was "slightly superior" to that of the monoglots.

All subjects in the monolingual and bilingual groups were given the Pintner Non-Language Intelligence Test and the Northumberland Standardized Test. The Northumberland Test, which is a verbal test of intelligence, was used chiefly as a check upon the non-verbal test.

The results indicated that the monolingual group had an average superiority of .8 of a year on the Northumberland Test whereas on the Pintner Non-Language Test, the bilingual group had an average superiority of .44 of a year. The boys in both language groups were found to be superior

to the girls on the non-verbal tests. Barke suggested that part of this inferiority of the girls might have been due to lack of interest in the material.

In drawing his conclusions from the investigation, the author advocated the use of a non-verbal test of intelligence when testing bilingual subjects in the following terms:

All that can be claimed is that we have here an indication that bilingual children will not prove inferior to monoglots (with a similar social environment) in an improved intelligence test from which the linguistic element is excluded. . . . It is indicated that under conditions of bilingualism, intelligence tests of a non-verbal nature should be used in preference or in addition to those in which success is conditioned by linguistic ability (6, pp. 249-250).

Mitchell (67), in an investigation which sought to determine whether an intelligence test administered in English is a fair estimate of a child's *IQ* when the child thinks in a language other than English, studied 236 Spanish-speaking pupils in grades one through three in a public school in Minnesota. The Otis Group Intelligence Test was administered in English and in Spanish to all the subjects. The mean *IQ* which has been secured from the Spanish testing was greater in each of the three grades than the mean *IQ* which had been secured from the English testing. For the three grades combined, the mean *IQ* was 9.28 points higher when the test was administered in Spanish. No reliable sex differences were found and the mean scores in the separate parts of the tests in either language show no substantial variation in performance on any particular part. The author concluded that bilingual children work under a serious handicap especially in the lower grades of the elementary schools and that the difficulty appeared to be a general language handicap rather than one in any specific phase of intelligence which the tests measured.

Seidl (88) performed an experiment the chief purpose of which was to determine the effect of bilingualism on the estimates of intelligence which were obtained on verbal and non-verbal tests. The 1916 Revision of the Stanford-Binet Scale was used as the verbal test of intelligence and the Arthur Point Scale of Performance Tests was used as the non-verbal scale. Seidl selected 240 subjects, all of whom were American born but of Italian descent, from the public schools of the same general area of New York City. There were 120 monoglots and 120 bilinguals and each of the language groups contained 60 boys and 60 girls. The age range of the subjects was from nine years six months to eleven years five months and the two language groups were closely equated to age. A questionnaire was used to determine

the selection of the language groups on the basis of the extent to which the Italian language was used in conversation with the child and the extent to which he used it. The Goodenough Revision of the Barr-Taussig Scale was used to classify the parental occupations. Both language groups were found to be skewed definitely to the lower end of the Goodenough Scale; the median for the monolingual group was in the Skilled Labor classification and the median for the bilingual group was in the Semi-skilled Labor classification. This lack of person for person matching between bilinguals and monoglots with respect to socio-economic status should be taken into account when the results are studied.

A statistical analysis of the results showed that the monoglots were superior to the bilinguals on all verbal tests. The average *IQ* of the monoglots in the Stanford-Binet Scale was 96.25 as against the average *IQ* of 91.61 for the bilinguals. On the performance tests, however, the bilinguals were superior to the monoglots; the average *IQ* of the bilinguals on the Arthur Scale was 100.41 as against an average age *IQ* of 96.21 for the monoglots. From a study of these results, Seidl concluded that if the Arthur Scale is accepted as a valid measure of intelligence, the two groups were about equal in native intelligence and that the difference which was found in the verbal test, which favored the monoglots, could be assigned to the language handicap of the bilingual subjects. The author did not explain, however, how he would account for the higher scores which were made by the bilinguals on the Arthur Scale.

The girls in each language group exceeded the boys on the verbal scale while the boys exceeded the girls on the performance scale. Seidl concluded that the sex differences were great enough to have a definite influence on the findings if he had not counteracted the sex factor by having an equal number of boys and girls in each language group.

Havighurst and Hilkevitch (47) reported an investigation in which seven trained examiners administered a shortened form of the Arthur Point Performance Scale to 670 Indian children, aged 6 through 15 years, in several Indian tribes in the plains and the southwest of the United States. The examiners found that the Hopi subjects received scores definitely above the norms of the white children on whom the full Arthur Scale had been standardized and that most of the other Indian tribes were approximately at the norms of the white children. The authors concluded that Indian children do about as well as white children on performance tests of intelligence.

Thirty boys and girls of the Sioux tribe were given the Kuhlmann-Ander-

son Test as well as the Arthur Scale and the results were compared. The average *IQ* on the Kuhlmann-Anderson Test was 82.5 with a standard deviation of 13.5, whereas the average *IQ* on the Arthur Scale for this group was 102.8 with a standard deviation of 19.1. The product-moment coefficient of correlation between the two sets of scores was $.53 \pm .09$. These results led the authors to conclude that "a performance test of intelligence would be more valuable for educational placement and guidance of Indian children in the southwest than an intelligence test which requires much use of the English language" (47, p. 431).

It must be remembered when viewing the results of this study that results on a shortened form of the Arthur Scale were compared with the norms on the full Arthur Scale. The number of children in the Sioux tribe who were given verbal and performance tests of intelligence was very small and the degree of bilingualism was not noted.

Shotwell (90) compared the differences in the results on the Stanford-Binet Scale and on the Arthur Performance Scale for 80 Mexican and 80 American mentally retarded patients of comparable ages at Pacific Colony. The *IQ*'s on the Stanford-Binet Scale for the total number of subjects ranged from 50 to 79. The mean *IQ* on the Arthur Scale for the Mexicans was 83 whereas the mean *IQ* for the Americans on this scale was 69. This difference was statistically significant. While both groups received higher scores on the Arthur Scale than on the Binet Scale the Americans averaged only 5 points higher on the Arthur Scale while the Mexicans averaged 22 points higher on this scale.

Of the Mexican group, 87.5 per cent received *IQ*'s on the Arthur Scale which exceeded their Binet *IQ*'s by more than 5 points, while 27.5 per cent received Arthur *IQ*'s which exceeded their Binet *IQ*'s by 30 or more points. While no Mexican had an *IQ* on the Arthur Scale which was more than 5 points lower than his *IQ* on the Binet Scale, 22.5 per cent of the Americans had *IQ*'s on the Arthur Scale which were more than 5 points lower than the *IQ*'s they received on the Binet Scale.

Shotwell concluded that Mexicans and other racial groups are "inadequately measured and unduly penalized" when they are measured by verbal tests of intelligence which have been standardized on an American white population.

Darcy (25) studied the effect of bilingualism upon the measurement of the intelligence of children of preschool age when the factors of age, socioeconomic status, and sex were held constant. The subjects were 212 American children of Italian parentage between the ages of two years, six months

and four years, five months. Of this number, 127 subjects were selected from the children attending 10 nursery schools in the boroughs of Brooklyn and Manhattan in New York City and 85 subjects were selected from the siblings of children who, at the time of selection, were attending a public school in the borough of Brooklyn. Of the 212 subjects, 106 were selected for the monolingual group and 106 were classified as bilinguals. A rating scale was used to determine whether the child qualified for the monolingual group or for the bilingual group.

If the answers to the questions on the rating scale indicated that the child heard and spoke Italian at home always or most of the time, whereas he heard and spoke English outside the home always or most of the time, he met the requirements for the bilingual group as far as the language factor was concerned. If, on the other hand, the child heard and spoke only English at home, spoke only English outside the home and heard English outside the home always or almost always, he met the requirements for the monolingual group as far as the language factor was concerned.

Both the monolingual and the bilingual groups were divided into four age levels, each covering six months of chronological age and there was a person-for-person matching in each group at each age level as to sex and also as to socio-economic status as determined by paternal occupations. The Goodenough Revision of the Barr-Taussig Scale was used to classify the occupations of the fathers. The factors of age, sex, number, and socio-economic status were held constant in order to control the variables which would be most likely to influence intelligence test results.

Two individual intelligence scales were administered to each child who was selected for this study. The 1937 Revision of the Stanford-Binet Scale, Form *L*, was used as the verbal test and the Atkins Object-Fitting Test, Form *A*, was used as the non-verbal means of measurement. The investigator assumed that if the performance of the bilingual subjects on the verbal test of intelligence were handicapped by the language factor, the individuals so handicapped would make higher scores on the non-verbal test. Conversely, if a significant difference were not found to exist between the scores of the bilinguals on the verbal and non-verbal scales, it might be concluded that bilingualism is not to be considered as a handicap when measuring the performance of bilingual subjects on verbal tests of intelligence.

A statistical analysis of the results achieved warrant the following conclusions with respect to the particular population which was studied:

1. There was a statistically significant difference between the mean *IQ*'s achieved by the monolingual and by the bilingual subjects on the Stanford-

Binet Scale. The mean *IQ* of the 106 monoglots on this scale was 98.69 with a standard deviation of 11.47 whereas, the mean *IQ* of the 106 bilinguals on this scale was 90.85 with a standard deviation of 11.37. The differences in mean *IQ*'s on the Stanford-Binet Scale were also consistently in favor of the monoglots when the two language groups were divided according to age and sex as well as when the age groups and sexes were combined.

2. Conversely, when the difference in the mean *IQ*'s on the Atkins Object-Fitting Test were determined for both language groups, a statistically significant difference was found in favor of the bilingual group. The mean *IQ* of the 106 monoglots on the Atkins Test was 88.95 with a standard deviation of 13.17 whereas the bilinguals achieved a mean *IQ* of 97.50 with a standard deviation of 15.30. The differences in mean *IQ*'s were consistently in favor of the bilinguals when the two language groups were divided according to age and sex as well as when the age groups and sexes were combined.

3. These data were substantiated by the differences which were found between the mean mental ages achieved by the two language groups on the Stanford-Binet Scale and on the Atkins Test.

4. The coefficients of correlation which were found between the mean *IQ*'s of the two language groups on the Stanford-Binet Scale and on the Object-Fitting Test were too low to warrant the substitution of one scale for the other in clinical or educational practice. However, these coefficients of correlation were sufficiently reliable to warrant the conclusion that the two scales are measuring the same functions to a large extent.

5. Since the monolingual and the bilingual subjects of this investigation were closely matched as to number, sex, socio-economic status, and age within six-month intervals, and since the performance of the bilingual subjects was significantly inferior to that of the monolingual subjects on the Stanford-Binet Scale but significantly superior to the performance of the monolingual subjects in the Atkins Object-Fitting Test, it may be concluded that the bilingual subjects of this investigation suffered from a language handicap in their performance on the Stanford-Binet Scale.

Portenier (73) studied the results on the Ohio State Psychological Test, Form 21, the Henmon-Nelson Tests of Mental Ability, the Terman-McNemar Mental Ability Test, and the Iowa Silent Reading Test which were administered to seniors who were enrolled in the Heart Mountain Relocation Project near Cody, Wyoming, in 1943, 1944, and 1945 respectively. Less than one per cent of these seniors were foreign born Japanese, 91 per cent were first generation Japanese-Americans, and the remaining 8 per cent

were second generation American-born Japanese. With respect to the bilingualism of the group, the author states that "since a large majority are first generation Japanese-Americans, the Japanese language may be spoken in their homes, in their social contacts, and more often in their religious life which is one of the more dominant influences perpetuating their culture" (73, p. 61).

There were about 200 seniors in each year but since the distribution of scores, the measures of central tendency and of variability, the centile ratings and the size of the respective groups were very similar for the three years, the 1945 group only was used for comparison with the high school seniors of all the schools in Wyoming.

The median and quartile scores on the O.S.U. Test for the Heart Mountain seniors were found to be consistently lower than those for all seniors in the state of Wyoming. The median score for the Heart Mountain seniors corresponded to the 47th percentile on the high school senior norms for the O.S.U. Test while the median score for all the Wyoming high school seniors for the same year ranked at the 55th percentile.

These results on the O.S.U. Test were verified by the findings on the Henmon-Nelson Test and the Terman-McNemar Test. The median *IQ* on the Henmon-Nelson Test for the Heart Mountain seniors was 97.61 which is in the lowest quartile of the norms for this test. Only 25 per cent of these students were above average on the basis of the norms of the test. The median score on the Terman-McNemar Test is equivalent to a mental age of 14-10 and only 25 per cent of the group had a mental age of 17-2 which is approximately the average chronological age of high school seniors.

The results on the Iowa Silent Reading Test indicated that one-half of the Heart Mountain seniors had a reading level in the lowest 35th percentile and that the upper 25 per cent had a centile rating of only 64 or above on the Iowa Reading Test. Portenier stated that the reading inferiority may have served as a handicap to the students when they took the verbal tests of mental ability.

D. STUDIES IN WHICH BILINGUALISM WAS FOUND TO HAVE NO EFFECT UPON THE MEASUREMENT OF INTELLIGENCE

Bere (8) sought to determine whether the differences between the results on mental tests of children of foreign parentage and those of native parentage were to be explained by differences in mental capacity or differences in language opportunity. Ten-year-old boys in New York City schools

were classified, by means of a questionnaire, into five language groups ranging from those who spoke only a foreign language at home to those who spoke only English at home. Three nationality groups, namely Hebrew, Bohemian, and Italian, were studied and there were 100 subjects in each group. The socio-economic status of each subject was determined but neither the nationality nor the language groups were matched according to this factor. The intelligence tests which were administered were the Pintner-Patterson Performance Scale and the Stanford-Binet Scale.

With respect to the Stanford-Binet results, the groups ranked in the following order: Hebrew, Bohemian, and Italian; while the order of rank on the Pintner-Patterson Performance Scale was: Bohemian, Italian, and Hebrew. Bere suggested that this difference in rank in nationality groups on the scales indicated the importance of using a wide variety of tests in measuring the abilities of different nationality groups. She also suggested that the groups which were studied varied in the nature as well as the amount of mental ability. All three of the nationality groups fell below Terman's "middle 50 per cent" on the Stanford-Binet Scale and received median *IQ*'s below 100.

A rise in scores on both the Stanford-Binet Scale and the Pintner-Patterson Performance Scale was concomitant with the increased use of English in the home. However, a study of the test results of those subjects whose mental ages on the Pintner-Patterson Performance Scale were higher than their mental ages on the Stanford-Binet Scale, revealed that those who used a foreign language at home exclusively, had scores which did not differ significantly from the scores of the other subjects. Bere stated that these findings suggested that the two scales measure abilities which might be correlated in some subjects and not in others. Furthermore, she concluded that the zero correlation between the results on both intelligence tests and the length of residence in this country of the parents of the subjects, plus the rise in both Stanford-Binet and Pintner-Patterson results with the increased use of English in the home, did not suggest that a language handicap was responsible for the low *IQ*'s on the Stanford-Binet Scale but rather that individual differences in nature may determine the extent to which an individual may profit by his environment.

Although Bere admitted that the results could not be applied to the population at large, the lack of matching of the nationality groups as to socio-status may have invalidated her results even for the population of her investigation.

Feingold (33), in an investigation which sought to determine the relative

rank of racial groups of the first generation of immigrants, studied the scores made on a modified form of the Army Alpha Test by 567 high school students in one public high school of Hartford, Connecticut. The test was given to all freshmen and two years later, to all who had become juniors and to all seniors. It would be of little purpose, in this review, to list the average relative rank of each of the racial groups which were studied. It is sufficient to state that the average *IQ* of the Italian Freshmen was 97 as compared with an average *IQ* of 103 both for the Americans (native-born) and for the Jewish group. The average *IQ* of the Italian Seniors was 96 as compared to the American native-born Seniors' average *IQ* of 100 and the Jewish Seniors' average *IQ* of 96.

The author concluded that the mental differences among the American-reared descendants of foreign races, and likewise the mental differences between American-reared children of foreign races and the children of Anglo-Saxons, are too small to be significant. It must be noted however, that there was no attempt to control other variables when measuring the intelligence of the groups which were studied. In addition, the number of children in each racial group, with the exception of the American native-born which numbered 264 and the Jewish group which numbered 146, was far too small to warrant any significant conclusions.

Darsie (26) and five field workers administered the 1916 Revision of the Stanford-Binet Scale and the Army Beta Test to 686 Japanese children in California. The author stated that the subjects constituted a non-selected sampling of the American-born Japanese population in the state, between the ages of 10 and 15 years. When the strictly rural children were excluded, there were 500 children to whom English was the more familiar language and whose scores could be adequately compared with the American groups on which the norms for the two intelligence tests, which were administered, were based.

The mean *IQ* of the Japanese 10-, 11-, 12-, and 13-year-old children on the Stanford-Binet Scale was found to be 91 as against a mean *IQ* of 99 for the American children of the same ages. The Japanese inferiority was limited mainly to the definitely linguistic items of the Stanford-Binet Scale but relative to this finding, the author stated: "It must not be overlooked that the existence of a pronounced language handicap may itself be indicative of a lack of capacity to master the language adequately" (26, p. 84).

On the Army Beta Test it was found that the mean scores of the two racial groups of the 10- and 11-year-old subjects were practically identical though the Japanese variability was greater. At year 12, the Japanese subjects were

superior to the Americans in the Symbol-Digit and Number-Comparison tests and in the tests dealing with Space Relations. The Americans at the 12-year level were superior to the Japanese in the Picture Completion Test.

The author concluded that while the differences in general mental capacity were slight, the Japanese subjects were inferior to the children of American and Northern European parentage in mental processes involving memory and thinking based upon concrete visually presented situations of a non-verbal character and in mental processes involving acuity of visual perception and recall and tenacity of attention. No significant sex differences in mental capacity were found for either group.

Darsie stated that the majority of his cases were found in the less desirable residential sections of the cities and towns in California and this factor in itself might lead one to suspect the advisability of comparing the Japanese subjects with the American subjects on whom the tests were standardized.

Hirsch (50) conducted an extensive investigation of 5504 subjects, all but 521 of whom were attending the public schools of four Massachusetts mill centers. Fifteen nationalities were represented and all subjects, except those in the American-white and the Negro groups, were American-born but of foreign parentage. The subjects, who were in grades one through nine, ranged in age from five years six months to 18 years no months but 97 per cent of the ages fell between 6 and 18 years. The social-economic factor was controlled only to the extent that the American and foreign subjects were chosen from the same neighborhoods. Although the experimental data include results from mental and educational tests; information bearing upon the relation of vocation to the degree of intelligence and also some anthropomorphic data, this review of the investigation will be limited, for the most part, to that phase of the study which treats of the differences in the intelligence test results which were found to exist between the 1,030 American and the 350 Italian subjects.

The Pintner-Cunningham Primary Mental Test was administered to the subjects in the first grade; the Dearborn Group Test of Intelligence, Form A, was administered to the subjects in the second and third grades, and the Dearborn Group Test of Intelligence, Form C, was administered to the subjects who were in grades four through nine. The mean *IQ* of the Italians who took these tests was 85.8 with a standard deviation of 11.94. Most of the *IQ*'s of this group were distributed between 67 and 99. The mean *IQ* of the American-whites was 98.3 with a standard deviation of 15.87. The difference in mean *IQ*, therefore, between the American-whites and the Italians was 12.5 ± 0.5488 in favor of the Americans and the author

stated that this difference was found to be statistically significant. Only 1.1 per cent of the Italians, as against 8.4 per cent of the Americans, received *IQ*'s above 120. Furthermore, only 14 per cent of the Italians received *IQ*'s at or above the average American *IQ* of 98.3.

In studying the effect of the use of a second language upon the intelligence test results, the investigator stated that if unfamiliarity with the English language were the cause of the low average *IQ*'s of certain nationality groups, it could be reasonably expected that there would be a progressive lessening of the influence of language in the intelligence test scores in the higher grades when the subjects had been in an English-speaking school system for many years. However, the results showed that the increase in mean *IQ* in the upper grades, as compared to the increase in mean *IQ* in the lower grades for the children whose parents came from non-English speaking countries, was found to be no greater than the small increase in mean *IQ* in the upper grades which was found in the groups of American-born children of English-speaking parentage. Great differences in mean *IQ*'s were found among the non-English speaking nationality groups but similarly greater differences in mean *IQ*'s were found among the English-speaking nationality groups. Several of the non-English speaking groups were found to have been mentally superior to several of the English-speaking groups.

In his analysis of these results, Hirsch reached the following conclusion with respect to the language factor: "Summarizing, we must conclude from an analytical inspection of our groups by grades, that a language handicap in American-born children of foreign parentage is a negligible factor in accounting for the mental difference found to exist among the group" (50, p. 313).

The fact that Hirsch's nationality groups were not controlled with respect to number, sex, age, or degree of bilingualism, and only to a minor extent as to socio-economic status, must be taken into account when his results are being studied.

Lester's (61) study was concerned with the results of administering performance tests to foreign-born children and with the advisability of using such tests in the sectioning of children in school systems. A retarded group of 26 first grade children, 18 of whom came from homes where Polish was spoken always and eight of whom were English-speaking monoglots, were the subjects of the investigation.

The 1916 Revision of the Stanford-Binet Scale and several performance tests were administered to all the subjects. The performance tests included the Saguin Form Board, the Mare and Foal, the Knox Cube, the Healy

Form Board, *A*, the Kohs Block Design, and the Porteus Maze. These performance tests were administered as group tests and since the author found considerable difficulty in administering the mazes to the foreign group, she questioned the results.

The average Stanford-Binet *IQ* for the foreign group was 80.5 and that for the English-speaking groups was 81. The correlation coefficients between the average *MA*'s achieved by the foreign groups on the Performance Tests and on the Stanford-Binet Scale range from .28 to .60. The foreign group was superior to the English-speaking group on the Healy and Knox Cube tests. The investigator therefore concluded that some performance tests more closely resemble the Stanford-Binet Scale than do others and that in classifying the children of foreign parentage, it would be more practical to combine selected performance tests with the Stanford-Binet Scale. The small and unequal number of subjects in each language group and the administration of the performance tests as group tests, might serve to invalidate the results of this investigation.

Hill (49) conducted a study to determine the effect of bilingualism upon the measured intelligence of children of Italian parentage. Bilingual children of the first, third, and sixth grades in a New Jersey elementary school were matched with monolingual children with respect to chronological age, sex, mental age, *IQ*, and socio-economic status. Bilingualism was determined by means of a questionnaire of Italian language background, a test of comprehension of spoken Italian, three tests of Italian word meaning, and additional information supplied by the visiting teacher of the school.

On the first grade level, 36 monoglots were compared with 36 bilinguals and an analysis was made of the performance of the two groups on the basis of the results which were received on the language and non-language sub-tests of the Stanford-Binet Scale. It was found that while the monoglots were superior to the bilinguals on those tests "involving a moderate understanding of the English language," the bilinguals were found to be superior to the monoglots on tests involving "an extremely high degree of efficiency of understanding and use of the English language."

On the third grade level, 36 monoglots were matched with 36 bilinguals. The Haggerty Intelligence Test, Delta *I*, was the test which was used to match the groups as to mental age. The matched groups were then given the Pintner Non-Language Test, List 8 of the Morrison-McCall Spelling Scale, and the Wittmer Cylinder Test.

On the sixth grade level, 50 bilingual children who were paired with 50 monolingual children, were given the Otis Self-Administering Examination,

Form A, List 8 of the Morrison-McCall Spelling Scale, the Army Beta Group Examination and the Wittmer Cylinder Tests.

Hill found no reliable differences in scores on verbal, non-verbal, and performance tests between the Italian children who heard and spoke Italian at home and Italian children who heard and spoke English at home. He concluded therefore, that the effect of bilingualism on the measured intelligence of Italian children who hear and speak Italian at home, may be disregarded.

It must be noted that since the language groups had been matched with respect to mental age, reliable differences in intelligence test scores between the language groups, at each respective age level, could not be expected. The study is of interest, however, since no reliable differences were found between the results which were achieved on the verbal and non-verbal intelligence tests by the bilingual and monolingual groups.

Arthur (4) studied the scores which had been made on the Kuhlmann-Binet Scale by 186 children of the kindergarten, first and second grades in a school in a small city where 75 per cent of the population was said to be of foreign birth or parentage. These scores were compared with the scores made on the Kuhlmann-Anderson Scale by these same children when they reached the seventh or eighth grades. For the three groups combined, the difference between the median Kuhlmann-Binet *IQ* and the median Kuhlmann-Anderson *IQ* was 7.0 with a *PE* of 4.95. Arthur therefore concluded that the Kuhlmann-Binet Scale when given by an experienced psychologist to a kindergarten, first, or second grade child from a foreign speaking home, can yield a high degree of reliability and predictive value as measured by the Kuhlmann-Anderson Scale five to seven years later, provided that the child had had as much as a year in an English-speaking school environment before receiving the Kuhlmann-Binet Scale. There was also a greater tendency for scores to increase among the children from non-English speaking homes than among the children from more typically American homes. No attempt was made to match the American and "foreign" population in this study nor was the socio-economic factor taken into account in comparing the scores of the children from the English-speaking homes with those from the non-English speaking homes. The study has been reviewed, however, since it substantiated the use of verbal intelligence scales with children from non-English speaking homes, at least under the specific conditions which the author described.

Arsenian (3), with the aid of three trained assistants, essayed by means of controlled experimentation, to determine the relationship between bilin-

gualism and the measurement of mental development. He confined his investigation to the study of the intelligence and social background of bilingual children, between the ages of 9 and 14 years, who attended public schools in New York City. The two main experimental groups consisted of 1,152 American-born Italian and 1,196 American-born Jewish children. There were, in addition, about 300 children of mixed races who were studied, but the results of the investigation of this group do not differ essentially from the results of the study of the main experimental groups and therefore, need not be noted further.

Bilingualism was measured by means of the Hoffman Bilingual Schedule to which were added seven questions from the Simms Score Card in order to measure socio-economic status. Age-grade status was determined by comparing the age and grade location of the subjects with the modal age-grade status of children in all elementary schools in New York City. The relation was then found between bilingualism and age, age-grade status, sex, and socio-economic status for each of the racial groups which were studied and comparisons of the results were made for the two experimental groups. The relationships between intelligence as measured by the Pintner Non-Language Test and the Spearman Visual Perception Test, Part I, and age, sex, race, and socio-economic status, were also determined.

The children of the two main experimental groups were divided into low and high bilingual groups on the basis of the Hoffman Bilingual Schedule, at each age from 9 through 14 years and for each natio-racial group separately. These groups were equated as to age, socio-economic status, and sex. It was then determined whether or not significant differences existed between the mental development of low and high bilingual children in each natio-racial group.

A group of 38 monoglots were matched with a group of 38 bilinguals, person for person, as to race, sex, socio-economic status, and age in months, in order to determine whether or not there were any reliable differences in intelligence between them. The Pintner Non-Language Test and the Spearman Visual Perception Test, Part I, were used to measure the intelligence of these groups.

The investigator found that the extent of bilingualism did not vary significantly from age to age in the groups and ages which were studied. There was no significant difference between the bilingual background of boys and girls taken separately, in either of the main experimental groups. In the groups studied, there was a correlation coefficient of $-.20$ between bilingualism and socio-economic status; a correlation of 0 between bilingual

status and age-grade status for the Jewish group, and a correlation coefficient of 0.20 between bilingual status and age-grade status for the Italian group.

No significant sex differences in intelligence, as measured in this study, were found. The correlation between intelligence and socio-economic status, as measured in this study, was positive but slight varying between .038 and .140. This might have been due to the fact that there was a limited range in the socio-economic status of the groups studied.

Using the Pearson coefficient of correlation, Arsenian found practically no relationship between bilingualism and intelligence for the groups which were studied. When the socio-economic status was held constant by the partial correlation technique, there was no essential change. No reliable difference was found between the low and high bilingual groups as far as the measurement of mental ability was concerned.

When a group of 38 monoglots and a group of 38 bilinguals were matched, person for person, as to race, sex, socio-economic status, and age in months, no essential differences were found to exist between the two language groups as to intelligence or age-grade status.

Arsenian, therefore, concluded that bilingualism did not influence favorably or unfavorably the mental development of the children of ages 9 through 14 years who were studied in this investigation.

In evaluating Arsenian's conclusions, it should be remembered that the use of group intelligence tests, when intelligence is the variable to be measured in a study, may be criticized since it is generally recognized that individual intelligence tests are a far more valid measure of intelligence, as it is measured at the present time. Furthermore, the author admitted that the Spearman Visual Perception Test had not been standardized and the correlation found between this test and the Pintner Non-Language Test was $.61 \pm .174$.

It might be questioned whether Visual Perception and the six factors which are measured by the Pintner Non-Language Test are adequate measures of that which is generally defined as intelligence. Certainly most psychologists would insist that the verbal factor is an important one in measuring intelligence and therefore this study, or any study of its kind, should include, though not limit itself to a recognized intelligence test of a verbal character in order to give the whole picture.

The measurement of only 38 monoglots and 38 bilinguals for comparative results, could be questioned on the basis of the small number of subjects employed.

Finally, while the fact that the tests were administered by different

examiners might not have had any effect upon the results, the possibility of a variable having been introduced to affect the results, is to be considered.

Pintner and Arsenian (71), in an investigation of the relationship of bilingualism to verbal intelligence and to school adjustment, studied 459 Jewish children all of whom were native-born and came from one school in New York City. The subjects ranged in age from 10 years one month to 15 years and were in grades five and six at the time of the experiment. The neighborhood environment of the subjects was such that they were in a position to hear and use English and Yiddish interchangeably.

The bilingual status of the subjects was measured by the Hoffman Bilingual Schedule and their intelligence was measured by the Pintner Intelligence Test, Form A, Grades IV-VIII. The Pearson coefficient of correlation between bilingualism as measured by the Hoffman Bilingual Schedule and the raw scores on the Pintner Intelligence Test was $-.029$.

The mean *IQ*'s of the 20 per cent of high and low bilinguals at either end of the bilingual group were compared and the differences in the means were not statistically significant. These two groups of high and low bilinguals were also given the Pintner Non-Language Test and the differences in the mean *IQ*'s of the two groups on this test, while not statistically significant, are less than the differences in the mean *IQ*'s of the same groups on the Pintner Intelligence Test, which is a verbal test.

The investigators concluded that, in so far as the population which was studied and the tests which were used in this investigation were concerned, the relationship between intelligence and bilingualism was "practically zero." They added however, that "in other populations, particularly non-Jewish and in younger ones, it is possible that a higher negative relationship between bilingual status and verbal intelligence may exist" (71, p. 259).

Spoerl (100) conducted an investigation in order to determine the effect of bilingualism upon the measurement of intelligence and upon the academic achievement of selected groups of students at the American International College. A "survey group" of 69 bilinguals and an "intensive study group" of 32 bilinguals were matched with corresponding numbers of monolingual students with respect to sex, age, and intelligence as measured by the Henmon-Nelson Test of Mental Ability. In addition, the intensive study group was matched with the control group on the basis of socio-economic status. In order that a student be admitted to the bilingual group, he had to have learned the two languages at or before entrance into the first grade of elementary school.

The experimental and control groups were then given the 1937 Revision

of the Stanford-Binet Intelligence Examination, Form *I*, and the Purdue Placement Test in English. Records of the subjects' academic achievement throughout their school careers were also studied.

Spoerl found no significant differences between the performances of the bilinguists and monoglots on the 1937 Revision of the Stanford-Binet Intelligence Examination, Form *L*, and the Purdue Placement Test in English.

Spoerl found that no significant differences existed between the performances of the bilinguists and monoglots on the 1937 Revision of the Stanford-Binet Scale. A slight inferiority on five of the verbal items of the scale was shown by the bilingual students. However, this slight verbal handicap was not sufficient to affect significantly the mean scores of the bilinguists on the entire scale.

The bilinguists received higher scores than the monolinguals on the Purdue Placement Test in English and a study of the records of academic achievement revealed that the bilinguists had done consistently better academic work than the monolinguals at every year level above that of Junior High School. The investigator stated that a compensatory drive arising from a feeling of environmental insecurity may have been a contributing cause of the superiority of the bilinguists in academic achievement.

It must be noted in reviewing the results of this study of Spoerl, that the two language groups were matched as to the scores achieved on the Henmon-Nelson Test of Mental Ability and therefore, it might have been expected that the differences in their performances on the 1937 Revision of the Stanford-Binet would not be significant.

Havighurst, Gunther, and Pratt (46) studied the results on the Good-enough Draw-a-Man Test which was administered to 325 Indian children, aged 6 through 11 years, of several Indian tribes in the southwest of the United States. The Indian children received scores which were superior to those received by a group of 66 white children of a small midwestern city. The mean *IQ* of the white children was 101.2 whereas the mean *IQ*'s of the nine Indian groups ranged from 102 to 117. The Indian boys in the majority of the tribes made scores significantly higher than the scores of the Indian girls.

The authors concluded that while the validity of the Draw-a-Man Test as a test of general intelligence is not substantiated by this study, the hypothesis that this test is valid for one aspect of intelligence, the ability to form concepts based on observation, is supported by the cultural data of several Indian groups and that Indian children, especially Indian boys, should be

expected to do well on such a test which does not include any necessary use of language. The degree of bilingualism was not noted in the report of this study.

E. SUMMARY

In summarizing the findings of the studies which were reviewed in this article, it may be stated the investigators who concluded that bilingualism had a favorable effect on the measurement of intelligence, are in the minority. Davies and Hughes (28), whose subjects were Jewish bilinguals and monoglots, and Stark (101) whose subjects were Irish bilinguals and monoglots, found that the mean scores on verbal intelligence tests for the bilinguals of their respective investigations, were superior to the mean scores of the monoglots. However, the results of these studies may be questioned for the investigators made little attempt to control the variables such as socio-economic status and degrees of bilingualism, which may have affected their findings.

The general trend in the literature relating to the effect of bilingualism upon the measure of intelligence, has been toward the conclusion that bilinguals suffer from a language handicap when measured by verbal tests of intelligence. Saer (84) found that rural bilinguals in Wales were inferior on verbal intelligence tests to urban bilinguals and also to urban and rural monoglots. Pintner and Keller (72) reported significantly lower scores for the bilinguals of their investigation on a specially revised edition of the 1916 Stanford-Binet Scale. Brown (15) found that bilingualism presented the chief difficulty in testing the foreign born subjects of his investigation though pupils who had attended an American school for one or two years received no lower scores on the Stanford-Binet Scale in English than on a translation of the scale in their mother tongue. Colvin and Allen (22) found the Italian subjects of their investigation inferior to the American subjects in their performance on the Stanford-Binet Scale and on the National Intelligence Test. While Pintner (69) found no significant difference between the mean mental ages of the combined foreign groups and the American group of his investigation on the Pintner Non-Language Test, so small a percentage of the foreign groups reached the median mental age of the American group on the National Intelligence Test that he concluded that caution should be exercised when measuring the intelligence of foreign children solely by verbal tests of intelligence. Smith (92) found a group of bilinguals in South Wales inferior to a group of monoglots on tests in free composition, mutilated passage, word-building, and analogies. Graham's

(42) findings supported the conclusion that the Italians of her study were significantly inferior to the Americans in mean *IQ* on the Stanford-Binet Scale. Wang (105) found that the Chinese students of his investigation ranged 9 points below the native-born subjects on the Ohio State University Intelligence Test. The American subjects of Mead's (65) investigation made higher scores than the Italian subjects on the Otis Group Intelligence Test. The performance of the bilingual Indians of Jamieson and Sandiford's (53) investigation was superior to that of the monoglots on the Northumberland Intelligence Test, the Pintner-Cunningham Primary Mental Test, and the Pintner Non-Language Test. Yoshioka (110) found the scores of the bilingual Japanese students of his investigation far below the American and the Japanese norms on the National Intelligence Test. Pintner's (69) Italian bilingual subjects received significantly lower scores than the monolingual subjects on the Pintner-Cunningham Test. Sanchez (85) found that the average mental age of his Spanish bilinguists increased on successive verbal intelligence tests. He attributed this gain in mental age to the added language facility that had been achieved during the intervals between tests. Barke's (6) findings indicated that the monoglots of his investigation were superior to the bilinguists on the Northumberland Standardized Test, whereas, the bilinguists were superior to the monoglots on the Pintner Non-Language Test. Mitchell (67) also found that the Spanish bilinguists of his investigation worked under a serious handicap when taking verbal tests of intelligence in the lower grades of elementary school. The Italian bilingual subjects of Seidl's (88) investigation received significantly lower scores than the monoglots on the 1916 Revision of the Binet-Simon Scale, whereas the bilinguists were significantly superior to the monoglots on the Arthur Point Scale of Performance Tests. Havighurst and Hilkevitch (47) concluded that a performance test of intelligence would be more valuable than a verbal intelligence-test for the guidance and educational placement of Indian children in the southwestern part of the United States. Shotwell (90) found the Mexican subjects of his investigation to be "unduly penalized" when measured by verbal tests of intelligence which had been standardized on an American white population. The preschool bilingual subjects of Darcy's (25) investigation made lower scores than the monolingual subjects on the Stanford-Binet Scale and significantly higher scores than the monoglots on the Atkins Object-Fitting Test. Portenier (73) found the bilingual Japanese high school seniors of his study to be significantly below the norms on several verbal tests of intelligence.

While a few studies in the field have found that bilingualism did not

serve as any handicap when verbal tests were used to measure intelligence, the general findings have been that bilinguals are penalized when their intelligence is measured on verbal tests of intelligence but that there is no indication of the inferiority of bilingual subjects when their performance on non-language tests of intelligence is measured against that of monolingual subjects. The findings of Bere (8) led to the conclusion that the increased use of English in the home did not accompany a rise in scores on the Stanford-Binet Scale. Feingold (33) concluded as a result of his investigation, that the mental differences as measured by the Army Alpha Test, between the American-reared children of foreign races and the American reared children of Anglo-Saxons, are too small to be considered significant. While the Japanese students of Darsie's (26) investigation were inferior to American students on the Stanford-Binet Scale, the scores of the two language groups on the Army Beta Scale were found to be practically identical. Significant differences were found between the scores of the American-whites and the Italians of Hirsch's (50) investigation, when both groups were measured by the Pintner-Cunningham Primary Mental Test and the Dearborn Group Tests of Intelligence, Forms *A* and *C*. These differences were in favor of the American-whites. However, with respect to the other nationalities studied, he found as great a difference between the mean *IQ*'s of the "non-English-speaking nationality groups" as among the "English-speaking nationality groups." He concluded, therefore, that bilingualism was not the cause for the lowering of mean *IQ* in any non-English-speaking nationality group. Lester (61) found no significant difference between the performances of a group of Polish bilinguals and a group of Polish monoglots on the Stanford-Binet Scale or on a group of performance tests. Arthur (4) found verbal tests of intelligence reliable measures of intelligence for children from non-English-speaking homes under specified conditions. Hill (48) found no significant differences between the performance of Italian bilinguals and monoglots on the Pintner Non-Language Test, the Otis Self-Administering Examination, the Stanford-Binet Scale, and the Army Beta Group Examination. However, it must be noted that Hill matched his groups with respect to mental age on other verbal intelligence tests before he began his investigation proper. Arsenian (3) found no significant differences between the performances of high and low bilinguals or between monoglots and bilinguals on the Pintner Non-Language Test or on the Spearman Visual Perception Test. In comparing the results of high and low Jewish bilinguals, Pintner and Arsenian (71) found no significant differences between the scores received by the two groups on the

Pintner Intelligence Test or on the Pintner Non-Language Test. The bilingual and monolingual college students of Spoerl's (99) investigation did not differ significantly in their performance on the 1937 Revision of the Stanford-Binet Scale. However, it must be remembered that the investigator had previously matched his groups on the basis of scores achieved in the Henmon-Nelson Test of Mental Ability. Havighurst, Gunther, and Pratt (46) concluded that Indian children should be expected to do well in a test of intelligence such as the Goodenough Draw-a-Man Test which does not include any necessary use of language.

REFERENCES

1. ARLITT, A. H. On the need for caution in establishing race norms. *J. Appl. Psychol.*, 1921, **5**, 179-183.
2. ARSENIAN, S. Bilingualism and Mental Development. New York: Teach. Coll., Columbia Univ., 1937. Pp. 164.
3. ———. Bilingualism in the post-war world. *Psychol. Bull.*, 1945, **42**, 65-86.
4. ARTHUR, G. The predictive value of the Kuhlmann-Binet Scale for a partially Americanized school population. *J. Appl. Psychol.*, 1937, **21**, 359-364.
5. ATKINS, R. E. The Measurement of the Intelligence of Young Children by an Object-Fitting Test. Minneapolis: Univ. Minn. Press, 1931. Pp. 89.
6. BARKE, E. M. A study of the comparative intelligence of the children in certain bilingual and monoglot schools in South Wales. *Brit. J. Educ. Psychol.*, 1933, **3**, 237-250.
7. BAYLEY, N. Factors influencing growth of intelligence in young children. In *Thirty-Ninth Yearbook of the N. S. S. E.* Bloomington, Ill.: Pub. Sch. Pub. Co., 1940. (Pp. 49-79.)
8. BERE, M. A comparative study of the mental capacity of children of foreign parentage. New York: Teach. Coll., Columbia Univ., 1924. Pp. 105.
9. BHATIA, C. M. Intelligence and National Reconstruction. Bombay, India: Hind Kitabs, 1949. Pp. 93.
10. LE BILINGUISME ET L'EDUCATION. Travaux de la Conference Internationale tenue a Luxembourg, 1928. Geneve: Bureau International d'Education, 1929. Pp. 182.
11. BOAS, F. Race, Language, and Culture. New York: Macmillan, 1940. Pp. 667.
12. BOVET, P. Bilingualism and education. *New Era*, 1933, **14**, 161-163.
13. BRIGHAM, C. C. A study of American Intelligence. Princeton, N. J.: Princeton Univ. Press, 1923. Pp. 210.
14. ———. Intelligence tests of immigrant groups. *Psychol. Res.*, 1930, **37**, 158-165.
15. BROWN, F. A comparative study of the intelligence of Jewish and Scandinavian children. *J. Genet. Psychol.*, 1944, **64**, 67-92.
16. BROWN, G. L. Intelligence as related to nationality. *J. Educ. Res.*, 1922, **5**, 324-327.
17. CALDWELL, F. F., & MOWRY, M. D. The essay versus the objective examination as measures of the achievement of bilingual children. *J. Educ. Psychol.*, 1933, **24**, 695-702.
18. CALLARD, J. A. Bilingualism and the Preschool Child. In *World Fed. of Educ. Proc.* Dublin: 1933. (Pp. 56-58.)

19. CHAUNCEY, M. R. The relation of the home factor to achievement and intelligence test scores. *J. Educ. Res.*, 1929, **20**, 88-90.
20. CHICAGO UNIVERSITY COMMITTEE ON EDUCATION, TRAINING, AND RESEARCH ON RACE RELATIONS. Bulletin No. 1. Chicago: Univ. Chicago, 1948. Pp. 55.
21. COFFEY, H. S., & WELLMAN, B. L. The rôle of cultural status in intelligence changes of preschool children. *J. Exper. Educ.*, 1936, **5**, 191-202.
22. COLVIN, S. S., & ALLEN, R. D. Mental tests and linguistic ability. *J. Educ. Psychol.*, 1923, **14**, 1-20.
23. COVELLO, L. Language used in Italian families. *Atlantica*, 1934, **16**, 369-371.
24. CREWSDON, J. Speech in an imperfectly learned foreign language. *Brit. J. Psychol.*, 1941, **32**, 82-99.
25. DARCY, N. T. The effect of bilingualism upon the measurement of the intelligence of children of preschool age. *J. Educ. Psychol.*, 1946, **37**, 21-44.
26. DARSIE, M. L. The mental capacity of American-born Japanese children. *Comp. Psychol. Monog.*, 1926, **3**, 1-18.
27. DAVIES, D. T. Bilingualism in Wales. *Proc. 5th Biennial Conf. Dublin, Ireland: World Fed. Proc.*, 1933. (Pp. 39-51.)
28. DAVIES, J., & HUGHES, H. G. An investigation into the comparative intelligence and attainments of Jewish and Non-Jewish school children. *Brit. J. Psychol.*, 1927, **18**, 134-146.
29. DECROLY, O. Essai d'application Test de Ballard dans les ecoles Belges. *L'Ann. Psychol.*, 1926, **27**, 59-93.
30. DELMET, D. T. A study of the mental and scholastic abilities of Mexican children in the elementary school. *J. Juw. Res.*, 1930, **14**, 268-279.
31. DEMAREST, R. Differences in results on five standard tests administered to Anglo-American and Spanish-American seventh grade boys. *Amer. Psychol.*, 1946, **1**, 224.
32. DIZIK, Z. J., & LUBINA, E. S. The mother tongue in test examinations. *Soviet Psikhotakh*, 1932, **5**, 37-44.
33. FEINGOLD, C. A. Intelligence of first generation of immigrant groups. *J. Educ. & Psychol.*, 1924, **15**, 65-82.
34. FREEMAN, F. N. The interpretation of test results with especial reference to race comparisons. *J. Negro Educ.*, 1934, **3**, 519-522.
35. FRITZ, R. A., & RANKIN, N. R. The English handicap of junior high school pupils from foreign speaking homes and remedial suggestions. *J. Educ. Res.*, 1934, **27**, 412-421.
36. FUDUKA, T. A survey of the intelligence and the environment of school children. *Amer. J. Psychol.*, 1928, **19**, 31-40.
37. GARRETSON, O. K. A study of the causes of retardation among Mexican children. *J. Educ. Psychol.*, 1928, **19**, 31-40.
38. GARTH, T. R. Race psychology. New York: McGraw-Hill, 1931. Pp. 200.
39. GARTH, T. R., & SMITH, O. D. The performance of full blooded Indians on language and non-language intelligence tests. *J. Abn. & Soc. Psychol.*, 1937, **34**, 376-381.
40. GIARDINI, C., & ROOT, W. T. A comparison of Detroit first grade tests in Italian and English. *Psychol. Clin.*, 1923, **15**, 101-108.
41. GOODENOUGH, F. L. Racial differences in the intelligence of school children. *J. Exper. Psychol.*, 1926, **9**, 388-397.
42. GRAHAM, V. T. The intelligence of Italian and Jewish children in the Habit Clinic of the Massachusetts Division of Mental Hygiene. *J. Abn. & Soc. Psychol.*, 1925, **20**, 371-376.
43. ———. The intelligence of Chinese children in San Francisco. *J. Comp. Psychol.*, 1926, **6**, 43-71.

44. GRAY, L. H. *Foundations of Language*. New York: Macmillan, 1939. Pp. 545.
45. HALPERN, F. The intelligence of bilingual Jewish girls as measured by the Stanford-Binet and Pintner-Patterson Performance Tests. Unpublished M. A. Thesis. Columbia University, New York, 1933.
46. HAVIGHURST, R. J., GUNTHER, M. K., & PRATT, I. E. Environment and the Draw-a-Man Test. *J. Abn. & Soc. Psychol.*, 1946, **41**, 50-63.
47. HAVIGHURST, R. J., & HILKEVITCH, R. H. The intelligence of Indian children as measured by a performance scale. *J. Abn. & Soc. Psychol.*, 1944, **39**, 419-432.
48. HILL, H. S. Correlation between IQ's of bilinguals at different ages on different intelligence tests. *Sch. & Soc.*, 1936, **44**, 59-90.
49. ———. The effects of bilingualism on the measured intelligence of elementary school children of Italian parentage. *J. Exper. Educ.*, 1936, **5**, 75-79.
50. HIRSCH, N. D. A study of natio-racial mental differences. *Genet. Psychol. Monog.*, 1926, **1**, 231-407.
51. HOFFMAN, M. N. H. *The Measurement of Bilingual Background*. New York: Teach. Coll., Columbia Univ., 1934. Pp. 74.
52. HU, P. C. A study of the intelligence of Anglo-Chinese children. *Eugen. Rev.*, 1938, **30**, 109-116.
53. JAMIESON, E., & SANDIFORD, P. The mental capacity of Southern Ontario Indians. *J. Educ. Psychol.*, 1928, **19**, 313-328.
54. JONES, H. E., CONRAD, H. S., & BLANCHARD, M. B. Environmental handicaps in mental test performance. *Univ. Calif. Pub. in Psychol.*, 1932, **5**, 63-69.
55. KENYERES, A. Comment une petite Hongroise de sept ans apprend le Français. *Arch. de Psychol.*, 1938, **26**, 521-566.
56. KIRKPATRICK, C. *Intelligence and Immigration*. Baltimore: Williams & Wilkins, 1926. Pp. 127.
57. KLINEBERG, C. *Race Differences*. New York: Harper, 1935. Pp. 367.
58. KOCH, H. L., & SIMMONS, H. A study of the test performance of American, Mexican and Negro children. *Psychol. Monog.*, 1928, **35**, 1-116.
59. LEIGHTON, D., & KLUCKHAHN, C. *Children of the People*. Cambridge: Harvard Univ. Press, 1947. Pp. 277.
60. LEOPOLD, W. F. *Speech Development of a Bilingual Child*. Evanston: Northwestern Univ. Press, 1939. Pp. 188.
61. LESTER, O. F. Performance tests and foreign children. *J. Educ. Psychol.*, 1929, **20**, 303-309.
62. LOEVINGER, J. Intelligence as Related to Socio-Economic Factors. In *39th Yearbook of the N. S. S. E.*, Bloomington, Ill.: Pub. Sch. Pub., 1940. (Pp. 150-210.)
63. LUH, C. W., & WU, T. M. A comparative study of Chinese children on the Pintner Performance and the Binet tests. *J. Educ. Psychol.*, 1931, **2**, 403-408.
64. MCCARTHY, D. *The Language Development of the Preschool Child*. Minneapolis: Univ. Minn. Press, 1930. Pp. 174.
65. MEAD, M. Group intelligence and linguistic disability among Italian children. *Sch. & Soc.*, 1927, **25**, 465-468.
66. MARZI, A. L'importanza del linguaggio nello sviluppo della intelligenza. *Contrib. Lab. Psico. Milano*, 1944, **12**, 527-528.
67. MITCHELL, A. J. The effect of bilingualism in the measurement of intelligence. *Elem. Sch. J.*, 1937, **38**, 29-37.
68. NEFF, W. S. Socio-economic status and intelligence; a critical survey. *Psychol. Bull.*, 1938, **35**, 727-737.

69. PINTNER, R. Comparison of American and foreign children on intelligence tests. *J. Educ. Psychol.*, 1923, **14**, 292-295.
70. ———. The influence of language background on intelligence tests. *J. Soc. Psychol.*, 1932, **3**, 235-240.
71. PINTNER, R., & ARSENIAN, S. The relation of bilingualism to verbal intelligence and school adjustment. *J. Educ. Res.*, 1937, **31**, 255-263.
72. PINTNER, R., & KELLER, R. Intelligence of foreign children. *J. Educ. Psychol.*, 1922, **13**, 214-222.
73. PORTENIER, L. G. Abilities and interests of Japanese-American high school seniors. *J. Soc. Psychol.*, 1947, **25**, 53-61.
74. POWERS, F. F., & METZLER, M. Successful methods of teaching English to bilingual children in Seattle Public Schools. Wash., D. C.: U. S. Office of Educ., Pamphlet No. 76, 1937. Pp. 17.
75. RAUBICHECK, L. Psychology of multilingualism. *Volta Rev.*, 1934, **36**, 17-20.
76. REPORT OF NEW EDUCATIONAL FELLOWSHIP, SIXTH WORLD CONFERENCE IN NICE, 1932. London: New Education Fellowship, 1933. Pp. 365.
77. REPORT OF THE IMPERIAL EDUCATION CONFERENCE IN LONDON, 1911. London Wyman & Sons, 1911. Pp. 267.
78. RIGG, M. Some further data on the language handicap. *J. Educ. Psychol.*, 1928, **19**, 252-256.
79. RODRIGUEZ, A. Are bilingual children able to think in either language with equal facility and accuracy. *Dept. Elem. Sch. Prin.*, 1931, **10**, 98-101.
80. RONJAT, J. Le Development du Langage observe chez un Enfant Bilingue. Paris: Champion, 1913. Pp. 166.
81. RYAN, W. C., JR. Dublin meeting of World Federation of Education Associations: Problem of bilingualism. *Sch. & Soc.*, 1933, **38**, 353-354.
82. SAER, D. J. An inquiry into the effect of bilingualism upon the intelligence of young children. *J. Exper. Ped.*, 1922, **6**, 232-240 & 266-274.
83. ———. Effect of bilingualism on intelligence. *Brit. J. Psychol.*, 1923, **14**, 25-38.
84. SAER, D. J., SMITH, F., & HUGHES, J. The Bilingual Problem. Averstwyth: Univ. Coll. Wales, 1924. Pp. 112.
85. SANCHEZ, I. Scores of Spanish-speaking children on repeated tests. *Ped. Sem. & J. Genet. Psychol.*, 1932, **40**, 223-231.
86. SCHLIEBE, C. Schrifttumsschau zur psychologie der zwersprachigkeit. *Deutsch Arch. Landeaf. Volksforsch.*, 1939, **3**, 475-488.
87. SEAGO, D. K., & KOLDIN, T. S. A comparative study of the mental capacity of sixth grade Jewish and Italian children. *Sch. & Soc.*, 1925, **22**, 564-568.
88. SEIDL, J. C. The effect of bilingualism on the measurement of intelligence. Unpublished Ph.D. Thesis. New York: Fordham Univ., 1937.
89. SERETA, K. E. A comparative study of 100 Italian children at the six-year level. *Psychol. Clin.*, 1924, **16**, 216-231.
90. SHOTWELL, A. A. Arthur Performance ratings of Mexican and American high-grade mental defectives. *Amer. J. Ment. Def.*, 1945, **49**, 445-449.
91. SMAYERS, F. Bilingualism in Belgium. *Proc. 5th Bien. Conf. Dublin: World Fed. Educ. Proc.*, 1933. Pp. 38-39.
92. SMITH, F. Bilingualism and mental development. *Brit. J. Psych.*, 1923, **13**, 270-282.
93. SMITH, M. E. A study of five bilingual children from the same family. *Child Devel.*, 1931, **2**, 184-187.
94. ———. Bilingualism in Preschool Children. *Proc. 5th Bien. Conf. Dublin: World Fed. Educ. Proc.*, 1933. Pp. 55.

95. ———. A study of the speech of eight bilingual children of the same family. *Child. Devel.*, 1935, **6**, 19-25.
96. ———. Some light on the problem of bilingualism as found from a study of the progress in mastery of English among preschool children of Non-American ancestry in Hawaii. *Genet. Psychol. Monog.*, 1939, **21**, 121-284.
97. ———. The effect of bilingual background on college aptitude scores and grade point ratios earned by students at the University of Hawaii. *J. Educ. Psychol.*, 1942, **33**, 356-364.
98. ———. Measurement of the vocabularies of young bilingual children in both languages used. *J. Genet. Psychol.*, 1949, **74**, 305-310.
99. SPOERL, D. T. Bilinguality and emotional adjustment. *J. Abn. & Soc. Psychol.*, 1943, **34**, 35-57.
100. ———. The academic and verbal adjustment of college-age bilingual students. *J. Genet. Psychol.*, 1944, **64**, 139-157.
101. STARK, W. A. The effect of bilingual on general intelligence: an investigation carried out in certain Dublin primary schools. *Brit. J. Educ. Psychol.*, 1940, **10**, 78-79.
102. TERMAN, L. M. Intelligence and its measurement: a symposium. *J. Educ. Psychol.*, 1921, **12**, 127-131.
103. TIREMAN, L. S. School problems created by the homes of foreign speaking children. *Calif. J. Elem. Educ.*, 1940, **8**, 234-238.
104. WALTERS, F. C. Language handicap and Stanford Revision of Binet-Simon Tests. *J. Educ. Psychol.*, 1924, **15**, 276-284.
105. WANG, S. L. A demonstration of the language difficulty involved in comparing racial groups by means of verbal intelligence tests. *J. Appl. Psychol.*, 1926, **10**, 102-106.
106. WATTS, A. T. *The Language and Mental Development of Children*. London: Harrap, 1944. Pp. 354.
107. WEST, M. Bilingualism with special reference to Bengal. Report No. 13. Calcutta, India: Bureau of Education, 1926.
108. WILLIAMS, W. J. National bilingualism through school work. *Proc. 5th Bien. Conf.* Dublin, Ireland: World Fed. of Educ. Proc., 1933.
109. WRIGHT, D. C., & MANUEL, H. J. The language difficulty of Mexican children. *J. Genet. Psychol.*, 1929, **36**, 458-468.
110. YOSHIOKA, J. G. A study of bilingualism. *J. Genet. Psychol.*, 1929, **36**, 473-479.

8829 Fort Hamilton Parkway
Brooklyn 9, New York

SOME SPECIAL ABILITY TEST SCORES OF GIFTED CHILDREN*

Department of Education, Hunter College

FRANK T. WILSON

A. THE PROBLEM

The pupils of the Hunter College Elementary School have, for the past 10 years, been admitted on the basis of *IQ* scores in the upper one per cent of their age groups. The average *IQ* for the 1947-48 year, according to the entering Binets of the children in the year of their admittance, was about 151.

It was of interest to know something of their abilities in special areas and therefore, insofar as it was practicable within rather severe limitation of research possibilities, a variety of special abilities was measured by standardized instruments. The testing was limited to 11-year-old groups and took place during various years from 1947 to 1950.

The age range of the Hunter College Elementary School pupils is from 3 to 11 years. "Promotion" is chronological and enrichment in a regimen of informal and interest-motivated activities is the underlying approach to meeting their growth needs.

Findings from tests in five special ability areas given to children during their last year in the school are reported herewith. Validity of administration was high. Directions were carefully followed with ease and clarity and the response by the children was excellent in terms of effort and interest. They characteristically were coöperative, accepted the tests as challenges, exhibited poise, and worked with unhurried care and dispatch. Only occasionally did any individual show any resistance or lack of seriousness.

In connection with some of the analyses of these findings use has been made of results of the Wechsler-Bellevue Scale I (8), which was given to the children within a year of the time the special ability tests were administered. The Wechsler provides both a verbal and a performance *IQ* as well as a full scale score, which adds some interest to the study. As shown elsewhere, the Wechsler clearly taps other abilities as well as some of the same that enter into the Binet scores, on which the children were ad-

*Received in the Editorial Office on March 1, 1951.

mitted to the school, or they do so in different proportions, since they characteristically ran 20-25 points lower than the Binets and since correlation of the Binet and Wechsler-Bellevue scores has been found to be $.49 \pm .04$ for 153 cases of these children.

B. ART

The Meier Art Judgment Test was given to 59 children. This test is based upon "the general assumption—that aesthetic judgment is one of the most important, if not the most important, single factor in artistic competence" (5, p. 4). The task set the subject is to choose, between two almost identical versions, the better one, so keyed because therein an artistic "principle functions to make for a greater aesthetic value." The alternate version presents the same picture except that in it "the functioning of the principle has been impaired." The answer sheet on which the subject marks his judgment informs him "regarding what aspect of the composition change has affected some principle, but the principle is not named." "The subject is to respond to the separate versions as to a whole." The "materials have been devised on the basis of works of established merit." All are in black and white only.

The test was given to 40 boys and 49 girls of the sixth grade and results are presented in Tables 1 and 2.

TABLE 1
THE FREQUENCY DISTRIBUTIONS, MEDIAN, AND QUARTILE DEVIATIONS OF THE MEIER TEST

	Boys	Girls	Total
107.5	1		1
103.5	1	1	2
99.5		6	6
91.5	7	6	13
87.5	5	7	12
83.5	5	11	16
79.5	6	8	14
75.5	8	5	13
71.5	3	2	5
67.5	1	2	3
63.5		1	1
59.5	2		2
N	40	49	59
Mdn	83.5	85.9	85.0
Q	± 7.25	± 5.25	± 8.86
Q3			93.86
Norms for grades 7, 8, 9 (N = 1145)			
Mdn			88.3
Q3			95.5

TABLE 2
CORRELATIONS OF THE MEIER ART JUDGMENT TESTS AND THE W-B I IQ SCORES

	N	W-B		Full scale
		Verbal	Performance	
Boys	40	.41	.26	.36
Girls	49	.39	.34	.41
Total	89	*	.30	.39

*Not computed.

Table 1 indicates very similar median scores for boys and girls, but a considerably wider variability among the boys. Norms are not given in the manual for Grade 6. Norms for 1,445 junior high school pupils, Grades 7, 8, and 9 are provided, however, and are about 88.3 for the median score and 95.5 for the upper quartile of scores. The upper quartile score for the Hunter Grade 6 children, obtained by adding one quartile deviation, 8.86, to the median score of the group, 85.0, is 93.86.

This level in the norms is considered critical in that, the manual says, "Individuals falling into this range should, other things being equal, find almost certain success in an art career." A large proportion of the upper quartile of these 6th grade children fell within the range of the upper quartile of the junior high school pupils on whom the norms were established. That fact seems to support the conclusion that there was a strong trend toward superior art judgment ability in this group of sixth grade children.

Table 2 is given to indicate whether or not the superior general ability of these children may have markedly affected their art scores. The coefficients for the total group of 89 children are presumably more indicative of trends than the coefficients for the smaller sex groups. These coefficients are fairly high and seem to indicate that general intelligence had perhaps some effect on art scores, but other factors were more largely determinant of art scores.

C. MUSIC

1. *Drake Musical Memory Test*

The Drake Musical Memory Test was devised to measure "capacity for musical achievement" (2, p. 2). "Discovery of innate talent is the chief aim" of the test according to the author. It is made up of 12 series of 2-bar melodies, increasing in difficulty and in span of memory, which are played on the piano. In each series the first melody is repeated one or more times with slight changes in either key, time or notes, or is played without change.

The pupils indicate for each repetition what change is made, or if it is the same. The score is the number of errors made.

Form *A* was given to several 11-year-old groups, a total of 144 children being tested.

TABLE 3
DISTRIBUTION, MEANS, AND *SD*'s OF DRAKE MUSICAL MEMORY ERROR SCORES, FORM *A*
(LARGE NUMBERS ARE LOW SCORES)

	Boys	Girls	Total
87-89	1		1
51-53	1		1
48-50		1	1
45-47			
42-44	1		1
39-41	1	2	3
36-38	2	2	4
33-35	8	4	12
30-32	4	8	12
27-29	9	6	15
24-26	12	6	18
21-23	14	7	21
18-20	8	13	21
15-17	6	10	16
12-14	3	5	8
9-11	5	4	9
6-8		1	1
<i>N</i>	75	69	144
<i>Mn</i>	25.18	22.42	24.07
<i>SD</i>	10.89	9.42	9.87
Norm Average			
Age Equivalent	22	22	

Table 3 shows scores by sex and for total boys and girls. Compared with the norms both the boys and the girls scored in the average age group of 22 years "based on 1,979 cases from the third grade through college," for Form *A*. By percentile rank for average 11-year-olds the group average is at the 87.5 percentile for boys and girls 11-13 years of age, averaging 12 years old.

In order to indicate whether or not general intelligence contributed markedly to this superior rating of musical memory ability, Table 4 is presented. It shows small negative coefficients throughout. Since the scores of the music test are error scores the relationship is an inverse one by definition. The table needs to be inversely interpreted, therefore, to express the relationships between the music abilities discriminated by the test and mental abilities measured by the Wechsler test. So interpreted, the relationships are in the main quite small in degree, and somewhat varied in

TABLE 4
CORRELATIONS OF THE DRAKE MUSIC MEMORY TEST ERROR SCORES AND THE W-B I
IQ SCORES

	N	Verbal	Performance	Full scale
Boys	69	-.03	-.27	-.20
Girls	61	-.27	-.01	-.22
Total	130	-.08	-.23	-.16

respect to parts of the test and to sex differences. The relatively low figures for both the number of cases and the size of the coefficients probably means that the differences obtained were of uncertain significance.

2. *Kwalwasser-Ruch Test of Musical Accomplishment for Grades IV-XII*

This test "is designed to measure the achievement of pupils in the typical public school music course" (3, p. 1), validated against "specifications adopted by the Music Supervisors' National Conference," 1921. Thus the content was not especially suited to use in this school where enrichment and adaptability of experiences are major principles determining young children's activities. The median scored by 45 children on this test of formal music knowledge, was 92.5, 3.5 points below that of the norm for Grade VI.

Variability was quite large, reflecting what was known to be true regarding learnings from private music lessons taken by many of the children. Boys seemed less accomplished on the whole than girls, although a boy made the highest score in the group.

D. SCIENCE

The elementary form of the Calvert Science Information Test was given to three 11-year-old groups in the spring of 1947. This test was standardized for use in Grades 4-6, but it was entirely inadequate to measure these children, since 59 of the 72 who took it scored above the top grade equivalent of 8.5+.

Accordingly, in the spring of 1949 Form A of the Ruch-Popenoe General Science Test, "primarily for Grades 8 and 9" was given to 19 boys and 23 girls of two 11-year groups. The test was "designed primarily to measure the accomplishment of pupils in general and elementary science courses in either the 8th or 9th grade" (6, p. 1).

The general reaction of the children to this test was unusual. Most of them were completely at a loss in regard to many of the exercises and omitted a great many altogether. There seemed to be an emotional element best described as frustration for many of the children who squirmed,

TABLE 5
DISTRIBUTIONS, MEDIANS, AND RANGE OF KWALWASSER-RUCH TEST OF MUSICAL ACCOMPLISHMENT

	Boys	Girls	Total
220	1		1
210		3	3
200		1	1
190		1	1
180	1		1
170			
160		1	1
150		2	2
140		2	2
130	1	2	3
120	3		3
110		3	3
100		1	1
90	1	1	2
80	1	1	2
70	1	3	4
60		4	4
50		1	1
40	1	2	3
30	1	4	5
20			
10			
0	3		3
N	14	31	45
Mdn	85	95	92.5
Range	1-221	31-215	1-221
Norm Grade VI			96

TABLE 6
DISTRIBUTIONS, MEDIANS, AND RANGE OF SCORES ON RUCH-POPENOE GENERAL SCIENCE TEST

	Boys	Girls	Total
35	3	1	4
30	2	1	3
25	8	3	11
20	2	6	8
15	3	8	11
10	1	4	5
N	19	23	43
Mdn	27.2	19.9	23.1
Range	12.5-37.5	12.5-37.5	12.5-37.5
Norms Mid-Year Gr. 8 or 9			28.1
Rho Science: W-B N = 41			.45

groaned, and mumbled about their situation. However, they made scores on both parts of the test and the medians for boys and for the total group of boys and girls were not far below the norms for the children who had had one semester of science in Grade 8 or 9. This is all the more significant in view of the fact that no special teacher of science has been provided in the school and what work the children had was under the direction of their regular classroom teacher with no special materials. The findings bear out opinions that gifted children tend to be especially interested in science and acquire unusual understanding in that area.

There were 41 cases for whom *W-B* scores were available. A rho correlation of .45 indicates that general mental ability had a fair degree of relationship to the science abilities, but probably did not have a marked effect on scores.

E. MECHANICAL ABILITY

The Stenquist Mechanical Aptitude Test II was given to 76 eleven-year-olds. Thirty-seven of the group were children in a special class of gifted children in a New York City public school whose *IQ* scores were similar in range and median to those of the *HCES* children. The Stenquist scores were also similar to those of the Hunter pupils and so in order to have a larger sampling for purposes of comparing sexes the two groups were combined.

Table 7 shows eight points difference in medians between the sexes—a finding consistent with well-substantiated fact that boys tend to surpass

TABLE 7
DISTRIBUTIONS, MEDIAN, AND QUANTILES OF STENQUIST MECHANICAL APTITUDE TEST II

	Boys	Girls	Total
60	1		1
55	4	1	5
50	5	1	6
45	4	6	10
40	5	11	16
35	3	5	8
30	3	12	15
25	3	6	9
20	1	3	4
15		2	2
<i>N</i>	29	47	76
<i>Mdn</i>	44	36	40
<i>Q</i>	8.67	6.36	7.59
Norm for 11-yr.-old children			28
Percentile Position			
of Medians	91	77	85
<i>T</i> Score Equiv.	± 1.2 <i>SD</i>	$+ .5$ <i>SD</i>	$+ .9$ <i>SD</i>

girls in achievement in mechanical ability. The variability was also greater for the boys. However, one girl scored at the interval next to the highest interval reached by the boys, and one boy scored at the interval just above that of the lowest girl.

The median score for each sex was much above the Stenquist norm for 1,087 children reported in 1922. The manual is not clear as to the sex of the standardization group as the words "children," "pupils," and "boys" are used in various connections. The score of one of the illustrative cases, a 14-year-old boy, was interpreted as being "exceeded by 97 per cent of boys of *his own age*" (7, p. 17). If the norms are for boys only then the 47 girls of this study presumably surpass the abilities of girls in general, even more than is shown by the comparison in the table.

The percentile positions and *T*-Scores used in the Stenquist norms show, of course, similar evidence of superiority of these pupils, the boys scoring at the 91 percentile and the girls at the 77 percentile of the standardization group, equivalent to about ± 1.2 *SD* and $\pm .5$ *SD* respectively, of the distribution of scores of that group.

In order to see what relationship to general intelligence might be indicated

TABLE 8
CORRELATIONS OF STENQUIST SCORES AND W-B FULL SCALE IQ*

Boys and Girls *N* = 39

Part I	.41 \pm .13	Part. III	.43 \pm .13
Part II	.51 \pm .12	Total	.57 \pm .11

*Only 39 of the children who took the Stenquist had also taken the W-B I.

correlations of the 39 *HCES* children and *W-B* scores is shown in Table 8 for the total Stenquist score and three component parts of it. All coefficients are fairly high, and even within the limits of their *PE*'s would indicate considerable, but far from major, overlap of abilities. Stenquist reported a Pearson *r* of .21 \pm .04, and that of 275 seventh and eighth grade boys in a NYC school who "were above average in general abstract intelligence 52 per cent were also above average in general mechanical aptitude." The standardization group was, presumably, more nearly representative of seventh and eighth grade boys, than the group herein reported.

The Detroit Mechanical Aptitude Examination, Form *A*, was also given to a group of 22 *HCES* boys and girls. The median score was 160.6, *Q* \pm 15, total range 109 to 202. According to Baker's Table 1, this is in the upper 8 per cent of scores for 11-year-old children, which includes scores 149-337. He suggests that *A* and *B* scores, according to "approximately 200 Detroit

counselors and teachers" indicate promise as "artist, civil engineer, dentist, laboratory teacher, osteopath, shop teacher, and surgeon" (1, p. 15)—an interesting selection, and names of occupations numerous represented in the parental vocations of parents of the *HCES* children.

The Revised Minnesota Paper Form Board Tests were given to 20 children in January, 1948. The 1948 manual says that the tests appear "to measure the ability to perceive spatial relationships" (4, p. 2). "Scores have predictive value for achievement in mechanical fields and shopwork. . . ." The test is designed for use in high schools, colleges, and for industrial purposes. Although for this reason it was very difficult for 11-year-old children it was administered to a small group of 11-year-old pupils of the school to see what these gifted children could do with it.

TABLE 9
DISTRIBUTIONS, MEDIANS, AND RANGES OF SCORES ON THE REVISED MINNESOTA FORM BOARD TEST

	Boys	Girls	Total
55		1	1
50			
45			
40	1		1
35	2	1	3
30	3		3
25	2	5	7
20	1	3	4
15	.1		1
<i>N</i>	10	10	20
<i>Mdn</i>	33.3	27.0	30.4
<i>Range</i>	16-41.2	20.6-55	16-55
<i>Mdn Norm Grade 9 & 10</i>			39

Table 9 shows considerable variability for both the boys and the girls, the range for both running from 16-55. A girl made the highest score and a boy the lowest one, although the boys' median was some six points above that of the girls. The norm on the group nearest these children in age was for boys and girls in Grades 9 and 10 who scored a median of 39. Compared with that older group the Hunter children scored a median of 30.4, equivalent to the 28th percentile for the 9th and 10th graders.

F. SUMMARY

Gifted 11-year-old children who had spent most of their school life in the Hunter College Elementary School for young gifted children showed, in general, superiority in abilities in art judgment, music memory, science,

and mechanical abilities, as measured by seven different standardized tests. The degree of superiority was most marked, apparently, in music memory, and least in music accomplishment as measured by the 1921 Kwalwasser-Ruch test.

The distribution of scores approached more closely to the normal curve in all of these measures, than did the *IQ* scores on which the children had been admitted to the school. Variability in the measures was also very marked.

Correlations of scores on these tests with recent Wechsler-Bellevue *IQ* scores of the children were from near zero to moderate in degree.

The findings corroborate other reports of trends toward excellence in abilities of various kinds among gifted children together with relatively unimpressive correlations with *IQ*. They also indicate pronounced individual variability among young gifted children in terms of their various individual abilities and in comparison with one another.

REFERENCES

1. BAKER, H. J., VOELKER, P. H., & CROCKETT, A. C. Detroit Mechanical Aptitudes Examination, Form *A*, Manual. Bloomington, Ill.: Public School Publishing.
2. DRAKE, R. M. Musical Memory Test, Forms *A* and *B*. A Test of Musical Talent, Manual. Bloomington, Ill.: Public School Publishing, 1934.
3. KWALWASSER, J., & RUCH, G. M. Kwalwasser-Ruch Test of Musical Accomplishment for Grades IV-XII, Manual. Iowa City: State Univ. Iowa, 1942.
4. LIKERT, R., & QUASHA, W. H. The Revised Minnesota Paper Form Board Test, Manual. New York: Psychological Corporation, 1948.
5. MEIER, N. C. The Meier Art Test: I. Art Judgment, Manual. Iowa City: State Univ. Iowa, 1942.
6. RUCH, G. M., & POPENOE, H. F. Ruch-Popenoe General Science Test, Manual. New York: World Book, 1926.
7. STENQUIST, J. L. Stenquist Mechanical Aptitude Tests, Manual. New York: World Book, 1922.
8. WECHSLER, D. The Measurement of Adult Intelligence. (Third Ed.) Baltimore: Williams & Wilkins, 1944.

Department of Education

Hunter College

695 Park Avenue

New York City 21

SOME SUGGESTIVE RESULTS IN PSYCHOMETRIC TESTING OF THE CEREBRAL PALSID WITH GESELL, BINET, AND WECHSLER SCALES*

Department of Psychology, University of New Mexico

DOROTHY J. SIEVERS AND RALPH D. NORMAN

A. THE PROBLEM

Psychological examination of patients with cerebral palsy is essentially a difficult problem, the difficulty resting in the fact that the disorder may affect both the motor and psychic apparatus of the individual. Thus, administration of either verbal or performance tests requires much skill, patience, and understanding on the part of the examiner as well as some modifications in testing itself. Cerebral palsy may bring on a degree of motor handicap which may vary in any given patient from very mild to quite severe involvements, the arms, legs, trunk, face, tongue, or any combinations of these organs being affected. In addition, visual, auditory, or other sensory handicaps may be present. Particularly important from the standpoint of the psychometric examination is the widespread extent of speech disorder among cerebral palsied cases, for from 60 to 80 per cent of those afflicted have speech defects (10).

Considering the effects upon intelligence, we find that the palsied conditions may be the etiological factor in actual mental deficiency. Contrary to older beliefs, however, only about 25 per cent of the cerebral palsied have the type of brain damage which produces mental deficiency, the other 75 per cent apparently falling within the range of normal intelligence (5). There are numerous reasons why individuals with this disability are wrongly labelled or diagnosed as mentally defective. Among these reasons are the following: First, they often are unable to respond well in standard testing situations, especially in the verbal sphere; second, their physical handicaps frequently limit their environment to the extent that they have little or no opportunity to learn, or at least to acquire information about the types of materials presented in psychological tests; third, their clumsy motor actions, facial grimaces, and speech defects sometimes give the appearance of mental deficiency.

Recent years of careful research in psychometric testing have emphasized

*Received in the Editorial Office on March 7, 1951.

the importance of personality factors in psychological test results. For example, Axline (2) has demonstrated how it is possible for children with behavior problems to be wrongly diagnosed as defective, with numbers of these improving in intelligence test ratings as a result of play therapy. Thus, not to be gainsaid either as a factor in test results is the effects of cerebral palsy on the personality of the patient. Wardley (14) has pointed out, for example, that fear is overdeveloped in spastic cases and in those with physical rigidity, whereas anger or frustration is not so easily aroused in spastics. In athetoids and tremors, there is more of an extroverted picture, although anger is overdeveloped in the former. Ataxic cases appear to have more normal emotions. Strauss and Lehtinen (11) have indicated that brain-injured children are often erratic, uncoordinated, uncontrolled, uninhibited, and socially unaccepted. Goldstein (9) has mentioned the "catastrophic" reaction in brain-damaged patients, a form of behavior resulting from the threat of the disorder; this reaction indicates the feelings of inadequacy as well as the inconsistency of the behavior patterns involved.

This paper is the result of an exploratory study using differing psychological tests with cerebral palsied patients of a wide age range. Since the number of cases involved is rather limited, the results are merely suggestive, pointing out trends which might be fruitfully explored further. The study reported here is not an experimental one, in the strict sense, but rather a clinical evaluation to indicate some of the problems encountered with several standardized psychometric instruments. We shall attempt to analyze the subtests of the scales employed to discover whether there is any significance attached to the various tests passed or failed. We shall also indicate what we believe to be significant factors in failures, as well as some important points regarding administration, values, and handicaps of each scale.

B. SUBJECTS

The subjects were 34 in number. Twenty-five were patients in the Cerebral Palsy Outpatient Clinic in Albuquerque, New Mexico; six were enrolled in the Cerebral Palsy Day School in that city; three were undergoing speech therapy in connection with the Speech Department at the University of New Mexico. The age range of the subjects was from an infant of nine months to an adult of 54 years. Two of the subjects had prenatal injury, four had postnatal injury, and 28 had injuries established as being caused by birth. Twenty-three of the subjects were spastics—six hemiplegics, three paraplegics, one triplegic, and 13 quadriplegics; eight

were athetoids, two were ataxics, and one was rigid. There were 17 males and 17 females.

C. TESTS USED

For subjects between infancy and three years of age, the Gesell Developmental Scales (8) were used primarily. The interview technique was found most practical here, and the parents were asked about the child's behavior on the various parts of the scales. Then, whenever possible, the tests at the appropriate age zones were administered. No formal estimates of intellectual ability were gained from such results, but rather an evaluation of the development of the child in the areas of motor, adaptive, language, and personal-social behavior.

Between the ages of two and three, Form *M* of the Stanford-Binet was given whenever the child was physically able to take the test. Often, just some of the subtests were presented, the other subtests being eliminated because of the child's inability to take them. This procedure is in essential agreement with the views of Bice (3), Bosley (4), and Strother (12), all of whom protest against the indiscriminate use of the present standardized tests with cerebral palsied patients. Strother, for example, argues that one must first determine the motor and sensory capacities used in tasks for measuring intelligence, urging that tests which are not affected by the child's disability should be selected. This was done with our cases, and in addition, no *MA* or *IQ* was calculated with some of them. If the complete scale was not given, the Gesell Scales were used to add to the information obtained from the Binet.

Between the ages of three and about 14, Form *M* of the 1937 Stanford-Binet was the principal scale employed, although again, modifications had to be made in a few of the subtests. For example, the time limit was removed from the "Stringing Beads" subtest since it excessively penalized the child with poor motor coordination. The same thing was true of the "Sorting Buttons" subtest; also if the child was particularly handicapped in the upper extremities, the "Motor Coordination" test at level II-6 was eliminated. In some cases, the physical and mental development of the child appeared to be so retarded that the Gesell Scales were applicable even at these higher age levels although the scales do not go beyond three years.

After the age of 14 (except in one case, age 13) all subjects were given the Verbal Scale of the Wechsler-Bellevue. The exception in this group over 14 was unable to take the Wechsler because of her lack of speech; her physical handicap made the giving of any type of performance test also impossible.

To summarize, it might be said that a clinical approach was used, rather than one which adhered rigidly to standardized directions. The various tests were tried with one purpose in mind—to obtain as accurate an *estimate* of the patient's intelligence as possible. The general procedure was to observe the patient at the first clinic to which he came and to obtain all the information possible about him through case records or individual interviews with relatives, etc. Through this information it was decided what type of testing situation would be most suitable to the patient's physical and mental development.

D. RESULTS WITH THE THREE SCALES

1. *The Gesell Scales*

a. Results with the Gesell Scales. Table 1 presents the summary of data obtained with the Gesell Scales for 14 cases tested. It indicates sex, *CA*, diagnosis, time of injury, developmental age in the four areas of the Gesell, amount of retardation, and amount of variation in weeks between the minimum and maximum developmental ages on the various parts of the Gesell Scales.

There are several facts immediately apparent from Table 1. In no case does any child test at age on the Gesell, retardation being apparent in all areas. This retardation may vary from the *CA* by only two months as in Case 5 in Adaptive Behavior or it may vary by as high as 61 months as in Case 19 in Language Development.

An outstanding fact about these cerebral palsy cases is that 12 of the 14 used in this study indicate uneven development—that is, variations from minimum to maximum in the four behavior areas of the Gesell Scales. These variations range from four weeks to 50 weeks as indicated in the last column of Table 1. The mean variation for the 14 cases is 17.1 weeks. This great variation can be expected when one is dealing with the brain-injured because the injury may have greater effects in one area of development than another.

Another fact which may be determined from Table 1 is that there seems to be a tendency for the variation to become greater as the child becomes older. If we split the 14 cases into two halves of seven cases each, we will get seven cases below *CA* of 3 and seven cases at or above this *CA*. The mean variation for those below 3 is 12.4 weeks; for those of *CA* 3 or above, this variation mean is 21.9 weeks. The three cases that show no variation have only slight physical involvements so that the physical handicap may be

TABLE 1
RESULTS WITH 14 CASES ON THE GESELL SCALES

Case No.	Sex	CA	Diagnosis	Time of injury*	Motor dev.	Adap. beh.	Lang. dev.	Per-soc. beh.	Prob. retard.	*** Range
1	M	0-9	Spastic Quadriplegia	N	4 wks.	4 wks.	8 wks.	8 wks.	7-8 mos.	4 wks.
2	M	1-1	Athetosis	N	24 wks.	24 wks.	28 wks.	28 wks.	5-6 mos.	4 wks.
3	M	1-7	Spastic Quadriplegia	N	40 wks.	40 wks.	52 wks.	52 wks.	7-9 mos.	12 wks.
** 4	M	1-8	Spastic Quadriplegia	N	28 wks.	16 wks.	16 wks.	28 wks.	13-16 mos.	12 wks.
5	F	2-2	Spastic Hemiplegia	N	15-18 mos.	2 yrs.	21 mos.	15-18 mos.	2-11 mos.	33 wks.
6	M	2-3	Spastic Paraplegia	N	56 wks.	18 mos.	18 mos.	18 mos.	9-14 mos.	22 wks.
7	F	2-7	Spastic Paraplegia	N	2 yrs.	2 yrs.	2 yrs.	2 yrs.	7 mos.	0 wks.
8	M	3-0	Spastic Quadriplegia	N	30 mos.	2 yrs.	2 yrs.	2 yrs.	1 yr.	26 wks.
** 9	F	3-1	Spastic Quadriplegia	PN	32 wks.	28 wks.	28 wks.	28 wks.	29-30 mos.	4 wks.
11	M	3-6	Spastic Quadriplegia	N	3 yrs.	3 yrs.	3 yrs.	3 yrs.	6 mos.	0 wks.
12	M	3-8	Spastic Quadriplegia	N	4 wks.	28 wks.	40 wks.	28 wks.	37-40 mos.	36 wks.
13	F	3-8	Athetosis	N	28 wks.	28 wks.	28 wks.	16 wks.	37 mos.	12 wks.
**15	M	4-4	Spastic Quadriplegia	N	15 mos.	40 wks.	40 wks.	52 wks.	37-42 mos.	25 wks.
**19	F	5-8	Ataxia	AN	18 mos.	40 wks.	40 wks.	40 wks.	52-61 mos.	50 wks.
	F	5-8	Athetosis	N	18 mos.	40 wks.	28 wks.	40 wks.	52-61 mos.	50 wks.

*Time of Injury: AN = Ante-natal, N = Natal, PN = Post-natal.

**Probable mental deficiency.

***Range is the difference between Maximum and Minimum.

a factor in the amount of variation observed in other cases. It is also true that the retardation seems to be more severe among the older cases. Among the cases below *CA* of 3, we obtain a rough Developmental Quotient of 57; for the cases at or above *CA* 3, the rough *DQ* is 34.¹

Among the 14 cases, there were four (Cases 4, 9, 15, and 19) in which mental deficiency was suspected and so stated in the provisional diagnosis. In three of these four cases (Cases 9, 15, and 19) it was noted that the motor development was more advanced than the development in other areas. The relative development of these four cases is illustrated in Figure 1. The remainder of the 14 cases do not show this tendency of advanced motor development so that the data from the Gesell might therefore suggest that in cases of cerebral palsy with mental deficiency, there appears to be somewhat more rapid development in motor behavior. The intra-test variability data may also serve as indications of the fact that other areas of the brain are more severely damaged than the motor area. These data, it may be said again, are only suggestive.

b. Factors in retardation on the Gesell Scales. All the cerebral palsied children in this study show some degree of retardation on the Gesell. What are some of the factors involved in this retardation? The present writers believe that their data, both from the tests themselves and correlated social histories, demonstrate the following facts:

(1). *Physical handicap.* The physical handicap and other defects such as poor speech and vision are important factors in the retardation of these children. All of them show retardation in speech; two or three have known visual defects. Because of these defects, they are unable to explore or palpate the world as normal children do, and their environment is accordingly definitely restricted. Such restriction is bound to appear on a scale such as the Gesell.

(2). *Mental deficiency.* In at least four cases, the possibility of mental defect as etiological in the retardation is not to be overlooked. The rate of

¹The writers have deliberately not reported *DQ's* in Table 1 since they are in entire agreement with the following statement: "Matching child to schedule and aiming at a determination of maturity is simple enough when development is symmetrical in the four fields of behavior. . . . When growth is asymmetrical, however, or when there is a great deal of scattering even within one field, the problem is less simple, though not insoluble. It is in just these instances that the psychometrically-minded would like to calculate and arrive at a compromise value; it is in just these instances that we insist that a compromise value is meaningless. . . ." (7, p. 20). Rough mean *DQ's* are mentioned here solely for comparative purposes.

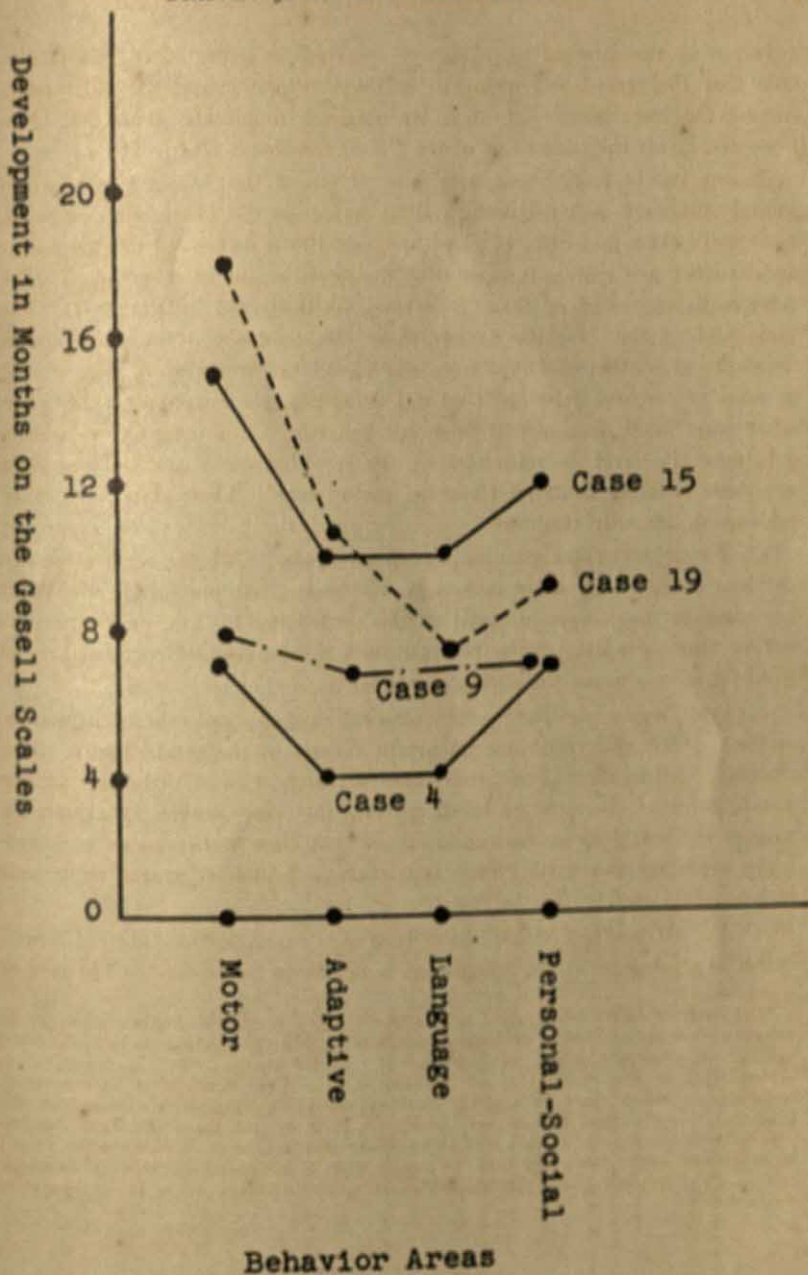


FIGURE 1
DEVELOPMENT OF FOUR CASES WITH POSSIBLE MENTAL DEFECT

development, of course, will be determined by the degree of mental defect, but it will in any case be much slower than the average.

(3). *Poor environment.* Inadequate environment appears to be an important aspect in at least five of the 14 cases. One comes from a common-law union where feeble-mindedness is suspected in the mother. Two cases come from homes where there is little or no comprehension about what to do with these handicapped children. One case has a psychotic mother; another has parents who are so disinterested that they will not make the effort to keep the child in the Day School. Certainly the type of environment is vital to every child, and these afflicted children are no exceptions. It may be said that their environment is of particular importance, on the other hand, since they must be given every possible opportunity to develop to the fullest extent.

(4). *Parental oversolicitation.* Parental oversolicitation appears to be a definite factor in at least four of the 14 cases; this is certainly a poor environment, but it is different perhaps from the type mentioned in the previous section in that emotional involvement with the child is probably deeper. The parents in these cases have let the children rule them; everything is done for the child, and his every wish is fulfilled. The children therefore have no motivation to learn to do things on their own.

(5). *Slow development.* Retardation is undoubtedly related to slow development in all of these cases. Doll *et al.* (6) earlier demonstrated a striking tendency toward delayed mental development among the cerebral palsied; they also found that early retardation was decreased in later life by prolonged development by smaller increments or by an increased rate of development. It may be anticipated then, according to Doll, that the cases in this study may show not as marked retardation in later years. This conclusion is substantiated by the fact that neither the Binet nor Wechsler *IQ*'s reported later in this study are, as low, on the average, as are the *DQ*'s on the Gesell.

2. *Revised Stanford-Binet Scale, Form M*

a. *Results with the Stanford-Binet Scale.* Table 2 presents the summary of data obtained with the Binet, indicating the sex, diagnosis, time of injury, *CA*, *MA*, *IQ*, and tests passed (indicated by x) or failed (indicated by -) on the various parts of the Binet. Twelve children were given the full scale; three (Cases 5, 7, and 8) were given only parts of the scale, being unable to take the full scale. Below, we shall discuss what appear to be significant aspects of failures on certain Binet subtests. However, it should

TABLE 2
RESULTS WITH 15 CASES ON THE REVISED STANFORD-BINET (FORM M)

Case number	5	7	8	10	11	14	16	17	18	20	21	22	24	25	26
Diagnosis	Spastic Hemi.	Spastic Para.	Spastic Quad.	Rigid	Spastic Quad.	Spastic Hemi.	Spastic Para.	Spastic Quad.	Spastic Quad.	Spastic Hemi.	Spastic Hemi.	Ataxic	Spastic Hemi.	Spastic Hemi.	Athetoid
Time of injury	N	N	N	N	N	N	N	N	N	PN	N	PN	AN	N	N
Sex	F	F	M	F	M	F	F	F	M	F	M	F	F	F	M
Chronological Age	2-2	2-7	3-0	3-3	4-0	3-9	4-11	5-0	5-7	6-6	7-0	7-3	9-7	12-8	14-7
Mental Age	-	-	-	3-1	2-10	3-1	5-0	2-8	4-6	4-7	4-5	5-4	6-6	9-0	6-6
IQ	-	-	-	95	71	82	102	53	81	71	68	74	68	71	46
Year II															
1. Delayed Response	x	x	x		x	x		x							
2. Ident. Objs. by Name	x	x	x		x	x		x							
3. Ident. Pts. of Body	-	x	x		x	x		x							
4. 3-Hole Form Board	x	x	x		x	x		x							
5. Picture Vocabulary	-	-	-		x	x		x							
6. Word Combinations	-	x	-		x	x		x							
Year II-6															
1. Ident. Objs. by Use			-	x	-	-		-							
2. Motor Coordination			x	x	x	x		x							
3. Naming Objects			-	x	x	x		-							
4. Picture Vocabulary			-	x	x	x		x							
5. Repeating 2 Digits			-	x	x	x		x							
6. Obeying Simple Commands			-	x	-	x		x							
Year III															
1. Block-Building: Bridge				-	-	-		-							
2. Picture Vocabulary				x	x	x		-							
3. Ident. Objs. by Use				-	-	-		-							
4. Drawing Vert. Line				x	-	x		x							
5. Naming Objects				x	x	-		-							
6. Repeating 3 Digits				x	x	-		-							

TABLE 2 (continued)

Case number	5	7	8	10	11	14	16	17	18	20	21	22	24	25	26
<i>Year III-6</i>															
1. Comparison of Balls				—	—	x		x		x					
2. Patience: Picture				—	x	x		—		x					
3. Discriminating An. Pic.				—	—	x		—		x					
4. Resp. to Picture I				x	—	—		—		x					
5. Sorting Buttons				x	x	—		—		x					
6. Comprehension I				x	—	—		x		x					
<i>Year IV</i>															
1. Picture Vocabulary				—	x	x	x	—	x	x	x	x			
2. Stringing Beads				—	—	x	x	x	x	x	x	x			
3. Opp. Anal. I				—	—	—	x	—	x	—	x	x			
4. Picture Identification				—	—	—	x	—	x	—	x	x			
5. No. Concept of 2				—	—	—	x	—	x	x	x	x			
6. Memory for Sent. I				—	—	—	x	—	x	x	x	x			
<i>Year IV-6</i>															
1. Discrim. Anim. Pict.					—	x	x	—	x	x	x	x			
2. Definitions					—	—	x	—	x	x	—	x			
3. Repeating 4 Digits					—	—	x	—	x	—	x	x			
4. Picture Comp.: Bird					—	—	—	—	—	—	—	—			
5. Materials					—	—	x	—	—	—	—	x			
6. Comprehension II					—	—	x	—	—	x	—	x			
<i>Year V</i>															
1. Picture Vocabulary						—	x		x	—	x	—			x
2. No. Concept of 3						—	x		—	—	x	x			x
3. Pict. Simil. & Diffs.						—	x		x	—	—	x			x
4. Patience: Rectangle						—	—		—	x	x	x			x
5. Comprehension II						—	x		—	x	—	x			x
6. Mutilated Pictures						—	x		x	—	—	x			x

TABLE 2 (continued)

Case number	5	7	8	10	11	14	16	17	18	20	21	22	24	25	26
Year VI															
1. Number Concepts							-		-	-	-	-	x		x
2. Copying Bead Chain							-		-	-	-	-	x		x
3. Differences							-		-	-	-	-	x		-
4. Resp. to Pict. I							x		-	-	-	x	x		x
5. Counting 13 Pennies							-		-	x	-	-	x		x
6. Opp. Anal. I							-		-	-	-	x	x		-
Year VII															
1. Giving No. Fingers							-			x		-	x	x	x
2. Memory for Sent. II							-			-		-	-	x	-
3. Picture Absurd I							-			-		-	-	x	x
4. Repeat. 3 Digits Rev.							-			-		-	-	x	-
5. Sentence Building I							-			-		-	-	x	-
6. Counting Taps							-			-		x	x	x	x
Year VIII															
1. Comprehension III										-		-	-	x	x
2. Simil: 2 Things										-		-	-	x	-
3. Verb. Absurd I										-		-	-	x	-
4. Naming Days of Week										-		-	-	x	x
5. Problem Situations										-		-	-	x	-
6. Opp. Anal. II										-		-	-	-	-
Year IX															
1. Memory for Designs I													-	x	-
2. Dissected Sent. I													-	-	-
3. Verbal Absurd II													-	x	-
4. Simil. & Diffs.													-	x	-
5. Rhymes: Old form													-	-	-
6. Repeat. 4 Digits Rev.													-	x	-

Notes: Case 25 passed two tests at Year X, and one at Year XI. Failing year was XII. Case 11 was 6 months older at time of Binet testing than when tested with the Gesell Scales

be borne in mind that the number of cases are few in number, so that our results are *suggestive*. Further research may or may not bear them out.²

b. Significant Failures on Stanford-Binet Subtests.

(1). *Identifying Objects by Use* (Years II-6, III). Four of the five children who took this test missed it; the fifth child passed it at Year II-6, but missed it at Year III. What is involved in this test? Terman and Merrill (13) state that it is probably language comprehension; they also mention that the objects chosen and their uses are familiar to children at this age level, but knowing an object differs from being able to identify it in response to a verbal description of it. We may now ask why this test appears to be quite difficult for the cerebral palsied child. On the Gesell it was seen that there was a general retardation in language development which includes language comprehension as well as actual speech ability. The one child who passed this test at II-6 had no speech defect, whereas others had defects of some degree, so speech ability may be one factor in failure. Still another factor could be the effects of the limited environment, a matter which has been discussed earlier in this paper. For instance, children who are unable to handle scissors are perhaps unfamiliar with them and thus fail in this particular item of the subtest.

(2). *Block-Building: Bridge* (Year III). None of the four children who took this test passed it. They had little or no conception of what they were to do here, but pushed the blocks around rather aimlessly. Lack of motor coördination did not seem to be a factor here because the test was attempted, and they all had good enough coördination to pick up the blocks. Then what are the factors responsible for failure? Gesell (8) lists this test among his tests for adaptive behavior, claiming that it depends upon motor skill and intellectual appreciation of form. Strauss (11) found that his brain-injured children suffered form-perception disability. Thus, it is postulated by the writers that poor form perception is an important factor in failure with this test among the cerebral palsied. [It is also apparent among brain-damaged adults, since Goldstein (9) speaks of the difficulty organic patients have in this respect.]

(3). *Picture Completion: Bird* (Year IV-6). Eight children took this test and not one of them passed it. Again, the lack of adequate motor coördination did not appear to be important in failure. Any attempt to

²The reader should also keep in mind that the accumulation of large numbers of cases of this disorder for study is no small task, and obtaining good test results reduces any potentially large *N*.

complete the bird would have been counted, but the most frequent response was merely a tracing around the lines already there, with the child generally considering this performance satisfactory. Terman and Merrill, interestingly enough, state as follows: "There is no evidence that the satisfactory completion of the picture is in any way related to drawing ability, even though success depends in part upon the development of motor coördination. . . . In the typical drawings at this level we can trace the characteristic stages of perceptual development of children . . ." (13, pp. 209-210). Once more, the possibility that cerebral palsied children have perceptual disturbances similar to those of the brain-injured children of Strauss arises as a factor in the failure with this test.

(4). *Materials* (Year IV-6). Only two of eight children passed this subtest. Terman and Merrill (13) claim that this is a test of language comprehension and information, with lack of success most often involving failure to comprehend the question. The same factors may therefore be operative in causing failure in this subtest as were mentioned in connection with Identifying Objects by Use.

(5). *Differences* (Year VI). Only one of the seven children who took this test passed it. Success in this test, according to its authors (13), is conditioned by several factors, the most important of which is the ability to make discriminating reactions; Terman and Merrill state that this ability is an early aspect of perceptual development. Other factors appear to be the degree of familiarity with the objects presented for comparison and the child's ability to keep a directing idea in mind. These cerebral palsied children again appear to be handicapped by perceptual difficulties and limited environment, both of which contribute to failure on this test.

(6). *Memory for Sentences II* (Year VII). One of six children passed this test. Shortness of attention span and poor language development taken together may be at the root of the difficulty. [Perhaps here we should say one in five failed, inasmuch as at least one case (Case 16), with a *CA* of 4-11 was working well above her *CA* level. This should be kept in mind also for comments on the next two tests mentioned below.]

(7). *Repeating 3 Digits Reversed* (Year VII). One of six children passed this test, which is primarily one of rote memory span. The shortness of attention span seemed to be the main cause of failure in this test, since the attention span increases very slowly in these children. This point will be discussed later.

(8). *Sentence Building I* (Year VII). Again but one in six chil-

dren passed this test, and it would appear that limited language development is at the core of failure. Some subjects did not know the meaning of the words they were to use to build sentences, and hence did not comprehend the nature of the whole task.

We have not discussed failures at levels higher than Year VII since many of the subjects were operating above their *CA* levels. Conclusions on failures, then, would be very tenuous.

Before leaving our discussion of the significant failures on Form *M*, it may be of interest to add a note here about three cases not included in this series which were examined with Form *L*.³ Case *A* was a female, *CA* 12-5, *MA* 3-9, *IQ* 30. Her basal age was Year II-6, failing age was Year VI. At Year III she failed Block-Building, Bridge, and Copying a Circle. At Year III-6, she failed Picture Vocabulary. At Year IV, failure was on Picture Vocabulary, Picture Completion: Man, and Discrimination of Forms. At IV-6, she failed Repeating 4 Digits and Materials. At V, she failed Picture Completion: Man, Paper Folding, and Copying a Square.

Case *B* was a male, *CA* 8-9, *MA* 5-2, *IQ* 59. Basal Year IV, Failure Year VIII. At IV-6, failed tests were Repeating 4 Digits and Materials. At Year V, failure was in Copying a Square and Memory for Sentences. Year VI indicated failures in Vocabulary, Copying Bead Chain, Number Concepts, and Pictorial Likenesses and Differences. At Year VII he passed only Repeating 5 Digits, surprisingly.

Case *C* was a male *CA* 2-10, *MA* 2-4½ (abbreviated Scale), *IQ* 85. Basal II, failure III. This child repeated 2 digits at II-6 but failed 3 at III. He failed identifying Parts of Body at II-6, also Block-Building: Bridge, Copying a Circle, and Stringing Beads at Year III.

Thus form perception again is a factor here, as may be attention span, in failure. The reader may note similarity in failed tests here with those in Form *M*, especially Block-Building, Materials, Memory for Sentences, and Repetition of Digits.

c. *Factors in poor performance on the Stanford-Binet.* If we consider only the *IQ*, only two (Cases 10 and 16) of the subjects tested with the Binet were in the normal range of intelligence. It is entirely probable that the same factors mentioned as causing retardation in the Gesell scales—physical handicap, mental defect, limited environment, parental oversolicitude, and slow development—play a rôle. It is impossible to say how many of these children are truly mentally defective—the numerical value of the *IQ*

³The authors wish to express appreciation to Miss Jane Powell of the Cerebral Palsy Day School for this testing, as well as for one Wechsler-Bellevue Test.

cannot be used as a reliable indication because these children cannot be compared with normal children who have, undoubtedly, different rates of development and a different degree of environmental stimulation.

Limited environment would definitely prevent these children from acquiring contact with the types of materials presented in the Binet, things that the normal child learns readily in his daily explorations. Such seems to us to be particularly true in the case of the test of Identifying Objects by Use.

An element which looms large in the Binet testing situation with these children is shortness of attention span among them. Everything was done during the testing to counteract its effect (giving test in more than one session, presenting tests quickly, removing distracting stimuli, etc.). Strauss (11) explains the short attention span in the brain-injured child by saying that he is abnormally responsive to stimuli in his environment. There seems to be a basic organic susceptibility to irritability—the environment, of course, being replete with irritating agents. The brain is involved in the phenomenon of attention, and damage may impair this function as well as others; attention is related to perception in that after attending to an aspect of the environment, clearer perception follows. This may be one phase of the perceptual difficulties of these children, namely, that the attentive focus is insufficiently delimited, causing failure on perceptual type of tests.

Limited speech is still another important element in causing failure on the Binet by these cerebral palsied children. The Binet, as is well-known, is quite verbal, and becomes even more so with ascending age levels. Difficulty with language comprehension appears to be a correlate of limited speech. One cannot, again, gainsay the effects of the environment in impoverishing the verbal accumulations of the palsied child, and these effects are several. In the first place, since the speech of the child is limited, parents and others may not talk to him as they do to normal children, so language comprehension in the child becomes retarded. Secondly, the child himself is incapable of participating in the linguistic world of his age-mates who could provide great stimulation in this important area of contact with the world.

d. Limitations of the Stanford-Binet. Several outstanding limitations of the Binet should be mentioned in connection with its use in testing cerebral palsied children. The first has been mentioned above—the verbal nature of the scale and the speech handicaps of these children. Along with this is the motor skill required on the test, but some of the subtests may be modified by removing the time limits and actually shortening the task such as reducing the number of beads to be strung in the Stringing Beads

subtest. This was the procedure used in this study. Of course, in modifying the test in this manner, the standardized testing situation is no longer extant, and this should be considered in interpretation of results such as ours. The standardized test situation is difficult to maintain, however, with most cerebral palsied children because one must perforce adapt the materials and situations to each individual's defects. The norms for the test are, of course, based on normal children in standard situations; these norms, however, cannot be used to evaluate the performance of these handicapped children because their environments have been entirely different and the testing must be altered.

It is our belief that the only legitimate use of the *MA* and *IQ* with these children is to contrast their successive performances with each other. Thus it is advocated here that *intra*-individual comparisons are of greater clinical value than *inter*-individual comparisons. This is also true of the Gesell Scales, or for that matter of any other scale—repeated examinations of the child will probably throw greater light on his development than comparing him with the norm.

In using the Binet with cerebral palsied children, careful attention must be paid to the disability of the child so that one may make plans beforehand for altering the testing situation. In the interpretation of results, it is imperative to scrutinize carefully each subtest for whatever light it casts on specific deficiencies in the intellectual sphere. The individual examination should be compared with previous tests, if any, to note if there are progressive, stationary, or deteriorative trends. The interpretation of results from testing should be correlated with findings from neurologist, orthopedist, speech correctionist, etc., as well as with social history data. Most important, a child with cerebral palsy should never be pronounced as mentally defective on the basis of a single administration of the Binet.

3. *The Wechsler-Bellevue Scale*

a. Results with the Wechsler-Bellevue Scale. Seven subjects were given only the Verbal Scale of the Wechsler since in at least six of the seven cases, the patients were unable to take the Performance Scale. In view of the limited number of cases here, all conclusions are again merely suggestive. Yet it is interesting to note that even with these seven cases there is material of interest. Table 3 summarizes the results with the Wechsler Scale, giving pertinent subtest information.

b. Significance of the Wechsler results. The order of success, then, of these cerebral palsied patients with the Wechsler-Bellevue is as follows, if

TABLE 3
RESULTS WITH SEVEN CASES ON THE WECHSLER-BELLEVUE (VERBAL SCALE ONLY)

Case No.	Sex	CA	Verbal IQ	Diagnosis	Time of injury	Mean W.T.S.	Deviations from subject's Mean W.T.S.					
							Info.	Comp	Digit span	Arith.	Simil.	Vocab.
28	F	19-0	90	Athetosis	N	7.8	+0.2	+2.2	-2.8	-1.8	+1.2	+0.2
29	M	26-5	85	Spastic Quad.	N	6.7	+1.3	+4.3	-2.7	-0.7	-1.7	+0.7
30	M	30-6	127	Spastic Quad.	N	13.8	+3.2	-2.8	-2.8	-3.8	+3.2	+3.2
31	F	31-2	101	Athetosis	N	9.2	-1.2	+1.8	-0.2	-0.2	-0.2	*
32	M	33-5	124	Spastic Quad.	N	12.8	+0.2	+3.2	-3.8	+2.2	-1.8	*
33	M	53-8	78	Spastic Quad.	PN	4.3	+1.7	+1.7	-1.3	-3.3	+0.7	+0.7
34	F	13-1	74	Athetosis	N	5.5	+0.5	-1.5	-1.5	+0.5	+2.5	-0.5
Mean Deviation from Subject's Mean Weighted Test Score							+0.8	+1.3	-2.2	-1.0	+0.56	+0.61

*These two subjects were not given vocabulary because of speech difficulty (Case 31) and time limitations (Case 32).

we study the mean deviations from the subject's mean weighted test score: Comprehension, Information, Vocabulary, Similarities, Arithmetic, and Digit Span. Each of these tests will be discussed briefly in the following sections.

(1). *Comprehension*. In general, these subjects performed best on this test. Four of the seven cases made their highest scores on this test and five of the seven have positive deviations from their *WTS* means. Wechsler (15) states that success on this test depends on practical information and the ability to use it or to evaluate past experience. Thus, even with the limitations which their handicaps impose upon them, these patients manage to collect sufficient practical experiences to pass the test. In the Stanford-Binet, it was noted, the children seemed to be hampered by lack of experience. We may conclude that the adult subjects have had a longer time to gather information by rather slow degrees.

(2). *Information*. This test is next in order of ease for the subjects reported on here. Six of the seven cases have positive deviations from *WTS* means. Wechsler (15) states that he tried to test the type of knowledge that the average individual has had an opportunity to acquire, so that possibly the remarks made about the Comprehension Test are pertinent here as well.

(3). *Vocabulary*. Only one subject, the youngest, had a negative deviation from his *WTS* mean, among the five given this test. Wechsler (15) and others have pointed out that this test is most resistive to organic deterioration. It appears to hold up comparatively well with these brain-damaged subjects.

(4). *Similarities*. Although the Mean Deviation is only a fraction of a point behind Vocabulary, there is more variability in this test, since three of the seven have negative deviations. This test taps the same type of ability as the similarities tests of the Binet. Previously, we stated that cerebral palsied subjects appear to have difficulties with similarities and differences because of perceptual disabilities. This factor may again be operative to make this test one of the more difficult ones in the scale.

(5). *Arithmetic*. The subjects, as a whole, did poorly with this test. The mean deviation is -1.0 , and five of seven have negative deviations. It is of interest here to note that Strauss (11) contends that sometimes the brain-injured person lacks the ability of the normal to learn the relationships of the number system spontaneously; this may again be related to perceptual disturbance. Perhaps another factor is lack of experience, since

at least three of the subjects had deficient educational backgrounds with possibly related lack of contact with arithmetic problems.

(6). *Digit Span*. This is by far the most difficult test in the Verbal Scale for these subjects. All seven cases had negative deviations, and the mean deviation of -2.2 is very significant. On the Stanford-Binet, it was found that children had great difficulty with this test, especially with digits reversed. Digits backwards gave these adult subjects a great deal of trouble. Once again, the matter of the deficient attention span seems to come to the fore. It is of no small significance, in our opinion, that even the two most intelligent subjects did not achieve more than average scores on this test.

Interestingly enough, our results are in very close agreement with those of Allen (1), who studied encephalopathic veterans on the Wechsler. The order of success for his brain-injured cases was (on the Verbal Scale): Information, Comprehension, Vocabulary, Similarities, Arithmetic, and Digit Span. Except for the inversion of order of Information and Comprehension, our results are in exact agreement with his. His Digit Span mean score had the highest critical ratio of difference of all subtests (including the Performance Scale) from either the Vocabulary subtest or the Information subtest. Regarding the Digit Span, he says, ". . . An important rôle is played by attention difficulties and the laboriousness of concentrated effort. While it is expected that the normal individual is not entirely free to receive external stimuli he does have greater receptivity than the brain-injured. An injury to the brain involves disturbances in receptivity, in maintenance of attention, and in the ability of immediate recall . . ." (1, p. 228). Our results may also be compared with Wechsler's "Hold" versus "Don't Hold" tests in his discussion of mental deterioration (15). Wechsler contends that the three "Hold" tests in the verbal scale are Vocabulary, Information, and Comprehension, whereas the three "Don't Hold" tests in this scale are Similarities, Arithmetic, and Digit Span. Thus, our three highest tests are the three "Hold Tests" and our three lowest are the three "Don't Hold." We did not calculate any "deterioration quotients" since we used only the verbal scale. However, we feel that it is quite striking that, with only seven cases, we are in agreement with Wechsler.

c. *Factors in performance on the Wechsler*. Factors discussed in this connection under the two previous tests used in this study apply here to a certain extent. The physical handicap of these patients is always an important consideration no matter what the age of the subject. With increasing age, the cerebral palsied may learn compensations for their handicaps as reflected by the generally better level of performance of these older cases.

(This is true, we believe, when we examine the Information and Comprehension Tests particularly.) There was no evidence of mental defect in any of the more mature cases, although poor environment still persisted, as was evidenced in case material obtained in studying them. At least three of the seven cases tested indicated inadequate environmental stimulation—Case 33 had no formal education, was unable to read or walk; Case 29 had devoted most of his spare time to listening to the radio for the past several years; Case 31 had spent much time living alone in a rooming house.

Generally, the Wechsler appears to give a more encouraging picture (keeping in mind, though, the limited number of cases) than does either the Gesell or the Binet. Two subjects are in the superior classification, one in the average, two in the dull-normal, and two in the borderline group, according to Wechsler's classification. The short attention span seems to have improved greatly, although it still is very poor. Limited speech was a factor which still persisted in these adults, especially in Case 33 where the subject seemed to be talking in short sentences of a few words each. But the other cases seemed to have adequate verbal facilities even though speech was unintelligible at times. They appeared to have overcome the retardation in language comprehension which was noticed at the lower levels.

An important question in test administration arises in relation to language—it is imperative that the examiner be quite familiar with the speech of these subjects before attempting to administer a verbal scale. If one has to ask them to repeat too often, rapport suffers inasmuch as such procedure engenders feelings of insecurity and uneasiness. In general, the urgency of applying the individual approach is mentioned again here, as it was in discussing the two previous scales.

E. SUMMARY AND CONCLUSIONS

This study is an investigation into the use of the Gesell, Binet, and Wechsler scales with 34 cerebral palsied patients ranging in age from nine months to 54 years, and presenting differing diagnoses. Each scale was generally used in the appropriate age group with certain exceptions, especially when for some reason of handicap the scale was inappropriate. Suggestive conclusions are as follows:

1. All of the cerebral palsied children used in this study show some degree of retardation on the Gesell Scales.
2. Children with cerebral palsy tend to show uneven development on the Gesell Scales in the four areas of behavior measured.

3. There is a tendency for older subjects tested with the Gesell to show greater variation among the four areas than younger cases.

4. There is some indication that mentally defective cases of the brain-injured type demonstrate most rapid development in motor behavior on the Gesell Scales.

5. Some of the causes for retardation on the Gesell Scales were intimated to be physical handicap, mental defect, poor environment, parental oversolicitation, and slow development.

6. On the Stanford-Binet (Form *M*), cerebral palsied subjects appear to have difficulty with the following subtests: Identifying Objects by Use at Year II-6; Block-Building at Year III; Picture Completion at Year IV-6; Materials at Year IV-6; Differences at Year VI; Repeating 3 Digits Reversed at Year VII; Memory for Sentences at Year VII; and Sentence Building at Year VII.

7. Factors in poor performance on the Binet were the same as those found influencing retardation on the Gesell. In addition there are shortness of attention span, perceptual difficulties, and limited speech, all three of which may be directly traceable to brain damage.

8. On the Wechsler-Bellevue adult subjects performed considerably better, on a comparative basis, than did children on either the Gesell or Binet. The order of success for adult subjects was Comprehension, Information, Vocabulary, Similarities, Arithmetic, and Digit Span. Only the verbal scale was given because these subjects could not handle the performance scale. Digit Span showed marked impairment as contrasted to other tests.

9. Order of success on the Wechsler agrees with previous work on the brain-injured. It corresponds with Wechsler's order of arrangement of "Hold" versus "Don't Hold" tests.

REFERENCES

1. ALLEN, R. M. The test performance of the brain injured. *J. Clin. Psychol.*, 1947, **3**, 225-230.
2. AXLINE, V. Mental deficiency—symptom or disease? *J. Consult. Psychol.* 1949, **13**, 313-327.
3. BICE, H. V. The psychological examination of the cerebral palsied. *J. Except. Child.*, 1948, **14**, 163-168.
4. BOSLEY, W. E. Objectionable objectivity in testing crippled children. *J. Except. Child.*, 1942, **8** (reprint).
5. BURTON, M. L. General implications of cerebral palsy. *J. Soc. Casework*, 1947, **28**, 91-95.
6. DOLL, E. A., PHELPS, W. M., & MELCHER, R. T. *Mental Deficiency Due to Birth Injuries*. New York: Macmillan, 1932.

7. GESELL, A. Booklet of General Instructions to Accompany the Gesell Developmental Schedules. New York: Psychol. Corp., 1940.
8. GESELL, A., & AMATRUDA, C. S. Developmental Diagnosis. New York: Hoeber, 1949.
9. GOLDSTEIN, K. The Organism. New York: American Book, 1939.
10. MILLER, M. M. Special education for the cerebral palsied child. *Crippled Child*, 1947, **25**, 4-5.
11. STRAUSS, A., & LEHTINEN, L. E. Psychopathology and Education of the Brain-Injured Child. New York: Grune & Stratton, 1947.
12. STROTHER, C. R. Evaluating intelligence of children handicapped by cerebral palsy. *Crippled Child*, 1945, **23**, 82-83.
13. TERMAN, L. M., & MERRILL, M. A. Measuring Intelligence. New York: Houghton-Mifflin, 1937.
14. WARDLEY, A. Psychotherapy and its implications in regard to cerebral palsy. *Spastic Rev.*, 1948, **9**, 5-6, 8.
15. WECHSLER, D. The Measurement of Adult Intelligence. Baltimore: Williams & Wilkins, 1944.

Department of Psychology
University of New Mexico
Albuquerque, New Mexico

PREJUDICE IN THE INTERVIEWS OF CHILDREN:
I. ATTITUDES TOWARD MINORITY GROUPS*

Institute of Child Welfare, University of California

ELSE FRENKEL-BRUNSWIK AND JOAN HAVEL

A. INTRODUCTION

This is a report on one of several aspects of a project on social discrimination in children¹ begun in the fall of 1944. During the course of this study approximately 1,500 children, most of them between the ages of 10 to 17, were tested. The tests² included scales on attitudes toward minority groups and toward political and economic issues, and a personality scale.

Of the children tested, 161 were selected on a basis to be described later, interviewed, and given the Thematic Apperception Test. The interviews concerned attitudes toward social and political issues, and such clinical data as attitude toward school, discipline, work, parents, friends, the opposite sex, and the image of the self. In addition, one or both of the parents of 43 of the children interviewed were also interviewed.

The present paper is the first of a series which will deal with the various sections of the interview. It is concerned with attitudes of members of the majority towards minority groups. Aside from presenting some of the relevant statistical evidence, its major purpose is the documentation of children's more or less spontaneous reactions to the minority problem, with summaries and interpretations in the manner of a running commentary.

A point that may be raised about this kind of study concerns the relation of verbal statements to overt behavior. Can the former be justifiably used as a basis for predicting how children will act in a concrete situation? While we have not included such an investigation in this study, we see no reason for a priori doubt about verbal behavior. Though we have evidence

*Accepted for publication by Harold E. Jones of the Editorial Board, and received in the Editorial Office on March 15, 1951.

¹This study was done under grants from the Research Department of the American Jewish Committee and the Rosenberg Foundation. A short preliminary report on the general outline and some first results of the project as a whole was given by the senior author of this paper some time ago (4). The manuscript of the present paper was read by Harold E. Jones, who made valuable suggestions.

²The tests were in the main devised by Else Frenkel-Brunswik, Murray E. Jarvik, and Milton Rokeach, who are authors of a forthcoming monograph on that aspect of the project.

from the interviews that certain children disguised their attitudes on the tests, we know from general clinical experience that non-verbal behavior is also subject to censorship. In a particular social setting a child may disguise dynamically potent hostility in action, though he might express it freely verbally, or in a changed situation. In another study (2) it was shown that short-sample behavioral observations are by no means superior to verbal statements as a basis for predicting long-range behavior.

Extension of the study of social beliefs to children has both advantages and handicaps. Children are generally more direct and uninhibited in their expression of attitudes and feelings than are adults. What children think about particular social issues has social implications, important to those concerned with changing attitudes. Children's attitudes, being less structured and less consistent than those of adults, are, to this extent, more susceptible to modification by environmental pressures. On the other hand, this very lack of structure and consistency presents difficulties for a study of attitudes.

The objection can be raised that, in studying the social beliefs of children, we are really studying those of their adult environment, of home and school. As far as the latter is concerned, we point to the wide range of attitudes we have found within a particular school, or even within a single classroom. This variation within an homogeneous school environment can, as will be seen, be attributed in large part to parental influence. We know from studies on adults (1), however, that prejudice is not an isolated phenomenon, but one of the expressions of a coherent, total pattern. Parents thus transmit to their children, not only specific social attitudes, but whatever predispositions to these attitudes lie within the individual's personality.

Considering the importance of the topic, the literature on the social and political outlook of children is relatively recent. Pioneer studies of the development of ingroup-outgroup images using scales, picture material, or a relatively spotty interview approach are those by Horowitz (6, 7), Minard (9), Zelig and Hendrickson (12, 13), and Peterson and Thurstone (10). Recently there is an increased interest in this topic expanding on and combining the previous methods (4, 11). On the whole there is considerable agreement as to results among the various investigators in this field. The present study is in many respects similar to that on adults mentioned above and presented in detail in a recent volume on the *Authoritarian Personality* (1). Unfortunately, the possibilities of direct comparison between the findings of this study on adults and the present one on children are limited by the fact that comparable materials for the adult subjects have not been evaluated quantitatively. However, there is a similarity between the present results and those obtained from adults by means of scales (1, 8).

B. THE SUBJECTS AND THEIR STANDING ON AN ETHNOCENTRISM SCALE

Only white American Gentiles were used as subjects for the interviews. Of the total of 161 children interviewed, the first 80 subjects were subjected to various tentative interviewing programs. The final interview schedule, developed on the basis of these preliminary studies, was administered to a further group of 81 children, 10 to 15 years of age. The material of the present paper is based on these later, complete interviews.

The 81 interviewees were drawn from three schools in Northern California. A total of 475 children had been tested for our project in these three schools, comprising about one-third of our total sample of 1,500.

The selection of our interviewees from the children tested was made on the basis of their score on one of the tests, a scale purporting to measure "ethnocentrism." The scale consisted of tolerant statements about Negroes, Mexicans, Japanese, Jews, and general outgroups, e.g., foreigners. Disagreement with these statements was taken as a measure of ethnic prejudice. The possible range of scores was from 0 to 100; the higher the score, the greater the prejudice as defined by the test. The designation high scorer, middle scorer, and low scorer, as used in this paper, refers to the total score made on the ethnocentrism scale, regardless of the specific minorities rejected and regardless of prejudice as expressed in the interview. The relationship between attitude-scale score and interview results will be presented later in this paper.

The first of the three schools selected, to be designated by the code-name, "Jackson," is a grade school located near the center of a large, cosmopolitan city. It is affiliated with a liberal college and is used for practice teaching. The school population is drawn from all over the city, and includes two particular kinds of children—those who are disciplinary problems in other schools, and those whose parents wish their children to attend a progressive school. The latter group of parents are often professionals, but the children in the school come from a variety of socio-economic backgrounds. All minority groups are represented in the school; of the 112 children tested at Jackson, 11 per cent were minority group members (about 4 per cent Jews, the remainder Negro, Mexican, Filipino, Chinese, and Japanese). The 34 children in the interview sample range from the 5th to the 8th grade; they were interviewed during the winter of 1946-1947.

"Marshall" is a grade school in a small, unincorporated community of 800 people, located a few miles from a large city. This community has developed since the last war, and contains a relatively homogeneous popula-

tion, primarily a lower-middle and upper-lower class group. Through the efforts of real estate promoters, the population has been restricted almost entirely to white Gentiles, though two Jewish children were enrolled in the school. Since the town is of recent development, however, many of the children have undoubtedly had previous contact with minority group members. Some of their parents, in fact, said that they moved to this neighborhood because they objected to the minority population in the area where they had been living. Thirty-one children, from the 5th to the 7th grades, were interviewed here, some in the spring of 1947, others early in 1948.

"Fairbanks" is a junior high school located in a town of 80,000 population. It draws from broad socio-economic strata and shows a varied ethnic and racial background. Approximately 16 per cent of the school population is Mexican, with a scattering of Negroes, Japanese, Chinese, and Jews. There are a large number of Portuguese and Italians, though these we did not consider minorities. Sixteen children, from the high 9th grade, were interviewed here during the spring of 1948.

The three schools differed in prejudice as measured by the ethnocentrism scale. The distribution of ethnocentrism-scale scores for Jackson was bimodal, with a mean of 49.0. That of Marshall was unimodal, but with a marked negative skew, and a mean of 61.9. Fairbanks came closest to a normal distribution of scores, though this too was negatively skewed, with a mean of 52.6. The quartile distributions relative to our total sample of 1,500 are shown in Table 1.

None of the three schools can in and by itself be considered representative of schools in this area. In fact, it is doubtful if any one school could be found that adequately represented children of this age. The schools were chosen primarily because of the willingness of principals and staff to give

TABLE 1
DISTRIBUTION OF ETHNOCENTRISM SCALE SCORES IN CHILDREN TESTED FOR THREE
SCHOOLS FROM WHICH INTERVIEWEES WERE SELECTED, RELATIVE TO TOTAL SAMPLE
FROM ALL SCHOOLS INVESTIGATED (NUMBER OF INTERVIEWEES
SPECIFIED IN PARENTHESES)

Range for total sample		Lower quartile 0-35	Middle quartiles 36-59 60-71		Upper quartile 72-100	Sums
Number of children tested (of these: interviewees)	Jackson	17 (7)	23 (5)	26 (10)	32 (12)	98 (34)
	Marshall	22 (14)	37 (3*)	32 (—)	61 (14)	152 (31)
	Fairbanks	52 (10)	74 (—)	50 (—)	49 (6)	225 (16)
	Sums	91 (31)	138 (8)	108 (10)	142 (32)	475 (81)

*These three cases have scores between 36 and 39 and are thus close to the low quartile.

time to this study; some schools refused participation because they feared pressure by parents or objections from minority-group members themselves.

Because of the complexity of the materials studied, the interview program concentrated on children scoring at the extremes of the ethnocentrism scale, in the hope of bringing out in their most pronounced form any differences in the response to interview questions related to different scores on the scale.

To counterbalance this one-sided approach, some children were also selected from the middle range of scorers. Though there were too few of these to make comparisons in the interview material, we felt it important, even within the limited framework of this study, to get some idea about the attitude of the majority of children, that is, those who fall between the extremes on this test.

The quantitative distribution of the interviews relative to the grouping of the children tested in the three schools and relative to the response tendencies of the total sample are shown in parenthetical figures in Table 1. There are slightly more than 30 interviewees in each of the lower and higher quartiles. The somewhat smaller group of intermediate cases interviewed was recruited mostly from Jackson; it is scattered fairly evenly over the middle half of scores.

For purposes of quotations in this paper, "high scorers" are defined as possessing ethnocentrism scores within the upper quartile of the distribution for the three schools selected rather than the distribution for the total sample (signifying a relatively high degree of prejudice); "low scorers" as being within the lower quartile, and "middle scorers" within the middle half. Since the sample in the three schools scores as a whole relatively high on ethnocentrism, only 27 (of the 32 upper quartile interviewees indicated on Table 1) qualify as "high scorers" in the present sense, whereas 36 (rather than 31) are to be considered as "low scorers"; the total of "middle scorers," 18, remains the same as in Table 1.

The preponderance of extremes among the interviewees limits our ability to generalize. However, in the present study we are concerned primarily with establishing certain pronounced patterns of ideology rather than with prediction from our sample to a larger group.

C. THE INTERVIEW SCHEDULE ON ATTITUDES TOWARD MINORITY GROUPS

Use of the interview technique permits us to compromise between the desire to obtain a maximum of spontaneous material and the necessity of

structuring the material sufficiently to permit some rough quantification and comparisons from child to child.

The program for the interviews is divided into several parts. To set a broad yet pointed frame, the issues raised at the beginning of the interviews dealt with social and political attitudes. Responses to the questions involved will be the subject of a separate paper.

The first of the items incorporated in the present report is the question, "What is an American?" The purpose of this question was to elicit whatever spontaneous mention of race or minorities a child might care to make without injecting the matter explicitly by a direct question.

This was followed by detailed questions on specific minority groups—Negroes, Mexicans, Chinese, Japanese, and Jews. The following questions were asked:

- What kind of people are_____?
- Why don't people like_____?
- What do people say about_____?
- What can_____do best?
- What could_____do to get along better?
- How do you feel about the segregation of_____?

There were also questions about attitudes toward minorities in such rôles as doctor, teacher, neighbor, president, friend, etc., and about acquaintance with minority group members.

This is a lengthy schedule, and time did not always permit a full discussion of all questions, or of all five minority groups. Considerable reliance was placed on the interviewer's discretion—when to pursue fruitful areas, when to avoid anxiety-producing ones. The questions asked and the sequence of questions were thus influenced to some extent by the nature of the child's response, the amount of material freely offered, his knowledge about and contact with minorities, his ease in the situation, and also by the limitations of time.

Last in the interview came the clinical aspects referred to in the Introduction. This material is also reserved for separate presentation in the future.

The children were told that the purpose of the whole interview was "to find out what boys and girls of your age think about different things"; it was not indicated at any point that the primary interest lay in minority group attitudes, although some of the children may have inferred this.

Sixty-five of the interviews were conducted by Mrs. Claire Bradner, who was unaware not only of the position of each child on the ethnocentrism

scale but also of the basic hypotheses concerning attitudes related to prejudice. The remaining 16 interviews were done by M. E. Jarvik, who was likewise unaware of the children's ethnocentrism scores but was familiar with the basic hypotheses of the study.

D. RATING AND CODING OF ATTITUDES TOWARDS MINORITY GROUPS: RELIABILITY AND VALIDITY

The evaluation of the interviews was "blind," that is, without any knowledge of the child's score on the attitude scale. The actual rating was performed by one of the present writers (Havel).

To secure some evidence of the reliability of this appraisal, the evaluation of 20 of the 81 interviews was independently duplicated by a second rater, Mr. Murray E. Jarvik.

Two major policies were adopted in evaluating the interviews. The first consisted in "over-all ratings" based on an intuitive synopsis of all the evidence relating to attitudes towards minorities, in the general manner proposed for the quantitative appraisal of qualitative material by one of the authors (Frenkel-Brunswik; for the execution of such ratings or scorings see 3, 1).

The second, more specific kind of rating—to be called "coding" here—was concerned with the presence or absence of some more narrowly circumscribed types of responses, such as the spontaneous bringing up of the topic of race in answering a particular question, or the rejection of a particular minority group if especially asked to make a statement, and the like.

Most important of the *over-all ratings* is a general estimate of prejudice. At this point we should define precisely the notions the raters were following when rating for prejudice. We did not assume that any negative statement made about a minority-group member, any reference to difference between a minority group and the majority is necessarily an expression of prejudice—though this is frequently, if not usually, the case. Nor did we consider a positive statement, in itself, an expression of tolerance. Prejudgment, in the true sense of the term, was taken to consist of the drawing of unwarranted conclusions on the basis of limited experience or of no experience at all. Thus, statements of difference between minority groups and the majority were considered as expressions of prejudice if and only if they were applied indiscriminately to all members of the minority group, or if they were spoken of as something natural and "given," or if they were made without attempt to explain them by reference to causal environmental factors. Unwillingness to grant minority-group members the same privileges extend-

ed to the majority, and undue hostility against a minority group were also taken as manifestations of prejudice. On the other hand, tolerance, perhaps better called absence of prejudice (not the absence of an attitude), was considered shown not only by an open stand against prejudiced behavior but also, and perhaps more frequently, by a willingness to accept members of a minority on the same basis as members of the majority group, by an assertion of the absence of fundamental or important difference between the two, or by the referral of existing difference to external pressures.

The overall ratings of prejudice were made on a seven-point scale graded in terms of what may best be called the "depth" of prejudice. The rating scale is as follows:

1. Positively and expressly tolerant; open condemnation of intolerant acts ("righteous indignation"); may go so far as to include a flaunting of one's own tolerance, though this is not essential for a "1" rating.
2. Tolerant—no prejudiced statements about outgroup(s) in question; like "1," but without the crusading spirit; no spontaneous condemnation (or very little of it) of intolerant acts.
3. Pseudo-tolerant: either a potentially unprejudiced person who is too passive to do anything but accept the prevailing spirit, or a person with intermediate attitudes who really doesn't care; tolerant statements generally prefaced by such remarks as "I've heard that" or "I don't know anything about them"; little involved in his attitudes.
4. Being the middle point, this can be almost anything; it may include a mixture of prejudiced and tolerant statements about a given outgroup; it may include admission of exceptions to one's own prejudiced statements, sometimes with a genuine feeling of puzzlement about these exceptions; or the attitude may be one that the individual feels is tolerant, though he basically considers the group in question inferior; may include intolerant statements prefaced by "I don't know" or "I've heard that. . ."
5. This rating to be given to a generally prejudiced person, but one who is not overly concerned about the problem, probably because of little contact; one who accepts his prejudices as a matter of course, who believes segregation is in general a good thing, but who is able to admit, perhaps grudgingly, that there may be exceptions to the rule; "All right if kept in their own place."
6. Open condemnation of out-groups who are considered without question to be inferior; no exceptions considered; will accept minority group members only at a distance, or in inferior positions, if at all.
7. Non-realistic condemnation and irrational fear of minority group members (as in the fascist).

There is obviously an interpretation process involved in making rating judgments of this kind. There also are undoubtedly halo effects from one

passage of the interview to another, strongly-worded prejudice or tolerant statements about some minority influencing the interpretation of the attitude as a whole. There is, moreover, the bias of the rater to consider, the bias being, as in the case of most investigators in this field, in the direction of condemnation of negative statements about minority-group members. The existing awareness of this bias, however, may have helped to minimize it.

The answer to the potential objections against intuitive ratings just listed is best given by strictly empirical considerations of reliability and validity.

As to reliability of the overall ratings, the correlation between the ratings on general prejudice made on a seven-point scale by the two raters independently for 20 of the children was computed and found to be .88. This result indicates that the obvious possibility of disagreement on the meaning of specific, often ambiguous statements taken in isolation is successfully overcome by using the total interview material in making the evaluation.

As to the validity of the overall ratings, the inferences incorporated in the overall prejudice ratings seem further warranted by the agreement of these judgments with an independent measure of prejudice, the overall score on the ethnocentrism scale. Before turning to quantitative evidence, we should like to insert a few general remarks on the question of the validity of our material. Although the ethnocentrism scale and the interview were intended as alternative approaches to an evaluation of the same attitudes, the differences between the two instruments are marked. The first attempts to measure prejudice by disagreement with statements made by someone other than the child; the second by the child's responses to relatively unstructured questions. The relationship between the child and the interviewer is without doubt very important, in some cases minimizing the expression of prejudice, because of a desire to please the interviewer, in others making its expression easier because the child feels safe and protected in the interview situation.

Considerable agreement between the two instruments exists in spite of these differences in approach, thus indicating that we are measuring something relatively consistent in the child. This agreement appears in two ways.

We shall first discuss differences in the overall interview ratings of prejudice just referred to when children scoring high and low on the ethnocentrism scale as a whole are compared with each other. In a computation of chi-square for ratings made for high and low scorers, the three rating values at the tolerant end of the scale were combined and the three at the prejudiced end of the scale were combined. Differences between the 27 high and the 36 low scorers, for both the tolerant and prejudiced ratings, were significant, in the expected direction, at the .001 level.

Secondly, the correlation between the interview rating on general prejudice and the total ethnocentrism-scale score was .67 for our definitive group of interviewees.

We have no evidence, as yet, of the greater validity of either instrument with respect to the overt behavioral aspects of prejudice. The fact that the results of the interview rating do correlate with the child's position on an ethnocentrism test adds support to the assumption that the interview tapped dynamically real attitudes in the child.

The *coding* as applied to the more specific and more directly palpable aspects of the interview was intended to summarize the child's statements, as much as possible in his own words. The difficulties in coding arose primarily from the nature of the interview itself—the variations in the questions asked, and in the amount of material elicited. Parts of the interview, e.g., those dealing with minority-group members in various rôles and with the child's acquaintance with minorities, had to be omitted from the quantitative analysis because of insufficient or ambiguous material.

A few brief comments must be added regarding the problems of reliability of the codings. From a common-sense point of view, the problem of the reliability of evaluation would seem to be of greater urgency in the case of the overall ratings than for the more specific codings, since the latter may more readily be defined in terms of the use of easily identifiable words or phrases by the child. Yet certain problems remain. Perusal of the entire material shows that existing discrepancies in coding between the two raters consisted largely of differences in the interpretation of when there was sufficient evidence to assume the presence of a certain kind of response rather than of differences in deciding which of several mutually exclusive alternatives was given in a particular instance. Percent agreements between the raters were computed for all items where there were sufficient data. With one minor exception, they range from about 75 per cent to close to 90 per cent. Procedures and data will be presented below in their proper places.

The interviews are meant in part to point up differences between children scoring high and low on the ethnocentrism scale and in part to characterize the attitudes of the children as a unified group. In the following pages the results will be presented from both standpoints—first a discussion of the responses of the group as a whole, then a more detailed description and concrete illustration of these responses, together with a consideration of the quantitative and qualitative differences between high and low scorers, and between boys and girls, where such differences are found to exist.

E. SPONTANEOUS MENTION OF RACE AND OTHER ATTRIBUTES IN DEFINING THE "AMERICAN"

As we mentioned above, the first interview question to be discussed in this paper, "What is an American?" was intended as a bridge between questions on political issues and those on race. It was made deliberately ambiguous—"American" is a concept that can be defined in terms of race or in terms of non-racial characteristics. Since race had not yet been mentioned by the interviewer, the children's answers give some indication of their readiness to respond in terms of racial categories.

The answers to this question were found to fall in one or more of a series of six categories of items. They are listed in Table 2. Each of the

TABLE 2
SPONTANEOUS MENTION OF RACE AND OTHER ATTRIBUTES IN ANSWERING THE QUESTION,
"WHAT IS AN AMERICAN?"

	Percent of answers containing spontaneous mention of category			
	Total group	High scorers	Middle scorers	Low scorers
(1) Race, national origin ^a				
(a) Condition for inclusion	11 47	14 33	28 62	— 51
(b) Denial of importance	36	19	34	51
(2) Good citizenship	40	37	39	43
(3) Conformity to law, custom	16	15	6	23
(4) Personal qualities	40	44	39	37
(5) Satisfaction with American way of life	5	4	11	3
(6) Material success, status	5	4	6	6

six items was coded for presence vs. absence in the answer of any given child by the rater (or raters) mentioned in the preceding section. "Undecided" was admitted as a third, intermediate coding judgment.

The reliability of this part of the coding procedure was established in terms of the agreement between the two raters on the 20 interviews coded by both of them. It was measured by the proportion of times they were in agreement on the items in Table 2. If one rater coded presence, the other absence of an item in a given case, this was counted as one disagreement; if one coded presence or absence and the other "undecided," this was counted half as agreement and half as disagreement. The percent agreement, defined as the number of agreements in the total of agreements plus disagreements, was found to be 80 per cent for the interview question under consideration.

The first of the items listed, that of nationality and race, is assumed to be present if the child refers directly or by implication to racial or national background in describing the American. The term "race" was found to have a variety of meanings for our children, depending both on their amount of information and their standing on the ethnocentrism scale, as will become apparent later from individual quotations and the description of differences between high and low scorers. General mention of "foreigners" has also been included with this category. As can further be seen from Table 2, it was found expedient to divide the first category into two subdivisions. Under (a) appear those subjects who explicitly or implicitly tend to establish national or racial criteria as a condition for inclusion under the concept of an American. Under (b) fall those with the opposite tendency, given by an explicit or implicit denial of the importance of nationality or race in this context.

As might be expected, these combined responses under the heading of nationality and race (Category 1a-b)—being almost literally suggested in the question—are the most frequent of all (47 per cent of the total group). Most of them (36 per cent) are in the negative, however.

A breakdown of the total group according to standing on the ethnocentrism scale is also given in Table 2. It is not surprising that this interview question, being rather general, proved to be less productive of differences between high and low scorers than those dealing directly with attitudes toward minority groups. The only category that differentiates between the two groups is the one now under consideration, the mention of nationality and race in the response.

Interestingly, it is the low scorers who most frequently mention the subject of race in their definition of an American (51 per cent of them, as opposed to a figure of 33 per cent for the high scorers). Although this difference is not statistically significant, it is large enough to warrant speculation as to its causes. High scorers may be more reluctant to raise the issue because of the setting in which they were interviewed. Or race, for them, may be a forbidden subject, to be discussed only when circumstances force a consideration of the question.

It must further be noted that low scorers, when they mention race, do so exclusively in the service of tolerance, that is, in order to affirm the right of anyone, regardless of birth or background, to be an American. Fourteen per cent of the high scorers, however, would exclude foreign-born or other outgroups from membership. The difference between the percent of the two groups willing to include anyone as American (19 per cent high

scorers as against 51 per cent low scorers) is significant at the 5 per cent level (as determined by chi-square, corrected for continuity). In both high and middle scorers the negative and the affirmative emphasis on nationality and race are almost evenly divided.

The answers in this category vary with respect to the size of the group embraced by the definition. Some children give all-inclusive answers—that is, they would admit anyone to membership as an American—sometimes adding a specific denial of the importance of race as a criterion for membership. Others make stringent requirements regarding birth and background.

The first kind of answer is most frequent among low scorers. It is illustrated by the quotations below:

B12, Age 10, Low.^a—Everybody who is in America. That is my opinion.

G20, Age 14, Low.—I think it's anything, practically, because America is made to be a country where every kind of person could come.

B33, Age 11, Low.—It doesn't have anything to do with race. Just being here makes you an American.

G23, Age 12, Middle.—Anyone can be an American. One of our teachers showed us a chart of all the races and said: "Pick out the typical American." And she taught us that we couldn't do such a thing, that everybody was an American.

B39, Age 15, High.—Should be honest, loyal, it doesn't matter what race or creed, should obey the laws.

The last quotation, from the interview of a high scorer, expresses the issue of tolerance in a more stereotyped manner, that is, in terms of slogans rather than in a direct, genuine manner.

One boy goes so far as to exclude the racial *in*-group from the definition, in a subtle ridicule of the bias implicit in the question.

B2, Age 12, Low.—No one is except an Indian. Being born here doesn't mean 100 per cent. Anybody is American that comes to America. Not all Indians are 100 per cent Americans. The ones born in these days aren't.

A more limiting definition, but perhaps a factually more correct one (because it refers to legal requirements) is given by others.

B29, Age 11, Low.—It is not necessary to be born in America, you can be a naturalized foreigner.

G19, Age 13, High.—Somebody who was born here in America.

^aEach quotation will be preceded by the following identifying data: Sex (B—boy; G—girl); Code number; Age; Position on the ethnocentrism scale (High scorer, Middle scorer, Low scorer).

Well, it doesn't have to be that—they can be from another country but believe the same things as the Americans who are here, have the same ideas. Or, if they get citizenship papers, they're Americans.

Only four of the children interviewed give openly prejudiced answers, specifically excluding foreign-born or particular ethnic groups. The high-scoring girl, quoted third below, goes even further and insists upon certain blood ties. Out of the four, only one boy—a middle scorer—shows ambiguity in his response and some inhibition in the expression of prejudice.

B21, Age 13, Middle.—One of our own nationality. He would be like that instead of a Chinese for example. You should go around with just your own race, your own color. My mother is a Czech.

G30, Age 14, High.—A person born here in America, only I wouldn't consider Mexicans and Negroes Americans even if they are born here.

G38, Age 10, High.—Well, he wouldn't have any German blood or any Japanese blood or any of that. He would be a *full* American.

B20, Age 13, Middle.—A white person—not only white but Negroes and stuff—say a white person.

A second category in the classification of responses deals with references to the American in terms of his functions and his obligations as a citizen. These include (according to the children) voting, loyalty to and support of the government, and fighting for the country in the event of war. As shown in Table 2, these qualities are mentioned by 40 per cent of the children. In this item, and particularly in those that follow, it appears that the question has been interpreted as: "What is a *good* American?"

One's obligations as a citizen are seen by some children as unquestioning submission to authority. This is exemplified by the following quotations from the records of two high-scoring boys, which seem to imply the request of abstention from criticism (in the first case, of the president, in the second, of Congress).

B4, Age 14, High.—Oh, he is loyal to his country. (?) Well, not to say bad things about the president; not plan to overthrow the country.

B38, Age 14, High.—Lives up to his country; don't tell secrets of his country to agents of Europe; agrees to all statements in Congress—don't put up a big argument about it.

The boy quoted next, a low scorer, does not emphasize criticism, but shows a general inhibition about taking action against authority.

B37, Age 15, Low.—Well, he's not necessarily born in America, as long as he is willing to fight for it, supports the government, doesn't do anything against America, tries to keep out of strikes and things like that.

Loyalty may take on this meaning of surrender to a personalized authority, but it may also be given a more obvious interpretation, as in the quotation below.

B19, Age 12, Low.—It means you have to be loyal to the country, respect the flag, be willing to fight in case of war.

G2, Age 13, Middle.—Being loyal to your country. We are all God's children—all good but brought up wrong—to swear and yell at brats. (She laughs).

The second category also includes answers, apparently more typical of the girls, referring to a belief in and adherence to American ideals. The two answers below are a good example of a real internalization of such beliefs.

G16, Age 13, Middle.—It means that we are equal, that everybody is as good as everybody else, and they have a right to live in America as well as we do. (?) They have to believe in what the U. S. does, freedom of speech and things like that. I suppose they don't *have* to, but they should.

G7, Age 14, Low.—To me it means he is born here or has citizenship papers, believes in America, is willing to accept its ideas and to vote. I think every American should vote. Some don't then fuss when things go the way they don't want.

Obedience to laws was considered as a separate, third category, since here the general concept of loyalty to the country is more depersonalized. The number of responses coded under this heading is relatively small, 16 per cent.

G5, Age 12, High.—One that obeys all the laws—is good to his country. (Explain?) I can't.

B43, Age 13, Low.—Means he is good, obeys all the laws, votes all the time, fights for his country if there is a war.

Note, in the next quotation, the denial of and reaction against aggressiveness, both in the fear of being different and in the reference to a generalized helpfulness.

G32, Age 14, Low.—One that should try to be a good citizen, do what other people do and not try to be different, follow the same laws that everyone else has to follow, try to do their best to help other people.

The last quotation is also representative of the fourth of the categories, reference to personal qualities. Some children will emphasize such external characteristics as health and cleanliness in defining an American. Others, however, stress more internalized values, such as honesty, unselfishness,

generosity, and helpfulness. The answers of 40 per cent of the children fell into this category.

The frequent vagueness of the internalized values is shown in the following quotations:

G17, Age 13, Low.—A person who is patriotic, who is a good sport.

G21, Age 13, Low.—Well, he does the best for his country. He always tries to help people out; trying to help out where he can and he always tries to do the right way.

B42, Age 10, High.—Well, he helps people. When people are poor he helps them and stuff like that.

In the third of the above quotations, that of a high-scoring boy, "help" has taken on a quality of condescension, being reserved for those considered lower in the social hierarchy.

Cleanliness, physical health and strength, and conventionalism—things known to be important to high scorers in the choice of personal friends—as conditions for being considered as American are brought out clearly in the answers of the two high-scoring girls quoted below.

G15, Age 14, High.—It's not always this way, but it's a clean sound. Physically, you think of somebody that is healthy, strong looking.

G24, Age 12, High.—Well, he goes to church and he keeps the city clean and goes by the rules.

Another high-scoring girl defines by citing the extreme example of an heroic death and sacrifice:

G35, Age 12, High.—Well, they could be talking about a man like he was in the war or something and the other man was going to get killed. He pushed him down and he got killed himself. That would be a real American.

The elaborations of personal qualities by the following low-scoring boy give an example of the less dramatic, more everyday kind of reasoning characteristic of this group.

B36, Age 14, Low.—Is always loyal to his country, has a suitable, honest job, interested in community service, interested in the growth of his children, so they may turn out to have the same ideals he has. He should be peace-loving, not out for power.

Considering sex differences, we find that personal qualities as defined by category (4) are mentioned by 50 per cent of the girls but only 32 per cent of the boys. Although the difference is not statistically significant, it lends support to commonly emphasized trends also found in other parts of our

interviews to the effect that girls, as a group, are more concerned with interpersonal relationships and with feelings than are boys.

A few further quotations may be of interest, not so much because they represent a frequent type, but rather because they document the glorification of the ingroup that forms such an integral part of the prejudice against minorities. The second of the quotations below is an excellent example of the stress on satisfaction with the American way of life (Category 5), and of such aspects of material success as "getting ahead" (Category 6).

B9, Age 11, Middle.—Is honest and born in America and likes America better than any other country.

B16, Age 13, Middle.—Is anybody that lives in America—not a dictator. Wants to go the way Americans go—get ahead like Americans. If they keep wanting to go back to the country they came from, they are not real Americans. They should like our land. They should have stayed in their old country if they liked it so well.

F. THE VARYING TARGETS OF ETHNIC PREJUDICE

We turn now to the main part of our results. Table 3 summarizes some of the most important findings of this paper, broken down in terms of the five minorities under consideration.

Taking the middle scorers as a representative core group of the total sample taking the Ethnocentrism test, from 41 per cent to 83 per cent express prejudiced attitudes with regard to the various racial groups. From this it may be a reasonable conjecture that the majority of children comparable to ours are on the prejudiced rather than tolerant side.

As can be seen from the parenthetical figures in the first column of Table 3, it is only the Negro who receives mention in all of the 81 interviews. As was pointed out above in describing the interview schedule, questions on certain minorities were omitted if time was short or there seemed little likelihood that the child in question could have had sufficient contact or information on them. Thus the data on Jews, Mexicans, Japanese, and Chinese are increasingly incomplete, in the order given, although even the last-named groups receive attention in more than half of the interviews.

The remaining figures in the first column of Table 3 present a classification of the responses in percent of the interviews containing any kind of mention of the particular minority. The seven-point rating scale used for the overall ratings of prejudice (see the preceding section) was here applied to each minority separately. As was the case above, "tolerant" again embraces ratings 1-3, "prejudiced" ratings 5-7; "unclassifiable" means either a rating of 4, or an irrelevant response, or a claim on the part of the sub-

TABLE 3
PERCENTAGE DISTRIBUTION OF INTERVIEWEES MAKING PREJUDICED AND TOLERANT
STATEMENTS ABOUT EACH OF FIVE MINORITIES

Minorities on which statements are made	Total group of interviewees	High scorers	Middle scorers	Low scorers	Significance High vs. Low*
<i>Negro</i>	(81)**	(27)	(18)	(36)	
Prejudiced (%)	60	96	83	22	.001
Tolerant (%)	30	4	11	58	.001
Unclassifiable (%)	10	—	6	19	.05
<i>Mexican</i>	(61)	(22)	(15)	(24)	
Prejudiced (%)	39	59	47	17	.01
Tolerant (%)	30	9	20	54	.01
Unclassifiable (%)	31	32	33	29	
<i>Chinese</i>	(43)	(14)	(12)	(17)	
Prejudiced (%)	44	57	67	18	
Tolerant (%)	21	7	8	41	
Unclassifiable (%)	35	36	25	41	
<i>Japanese</i>	(44)	(19)	(10)	(15)	
Prejudiced (%)	41	53	50	20	
Tolerant (%)	20	5	10	47	.05
Unclassifiable (%)	39	42	40	33	
<i>Jew</i>	(75)	(25)	(17)	(38)	
Prejudiced (%)	32	44	41	18	
Tolerant (%)	35	16	35	48	.05
Unclassifiable (%)	32	40	18	33	

*Level of confidence reached or bettered by the difference between high and low scorers is here given in terms of customarily used p -values .05 (5%), .01 (1%), .001 (one tenth of 1%).

**Italicized numbers in parentheses give the N-s on which percentages below are based, that is, the number of interviewees referring to the minority in question. Note that the figures beneath each of the parenthetic N-s add up to 100 per cent in each case.

ject to know nothing about the group in question. So far as the totals of responding interviewees are concerned, the highest per cent of prejudiced responses is for the Negro (60 per cent); the lowest is for the Jew (32 per cent); and the remaining three minorities are in the middle (about 40 per cent; the percentages based on the total number of children making statements about the minority in question, not on the total of all interviewees).

We turn now to differences in the expression of prejudice related to differences in ethnocentrism scale score. The major portion of Table 3 presents a breakdown of the results by high, middle, and low scorers. The differences between high and low scorers are all in the expected direction.

As seen from the last column of the table, they are statistically significant for the Negro and the Mexican. In the case of the Chinese and Japanese the differences are somewhat similar to those obtained for Mexicans; because of the small number of cases involved, most of them are not statistically significant, however. The difference between the percent of prejudiced statements made by high and low scorers about Jews is smaller than for any other group (44 per cent for high scorers vs. 18 per cent for low scorers) and likewise not statistically significant. Nevertheless, the direction of all the differences is quite consistent.

Several further trends appear in Table 3. Almost all the high scorers express prejudice against the Negro, but only approximately half of this group makes prejudiced statements about other minorities. This does not mean, as an inspection of the table will reveal, that high scorers make tolerant statements about other minority groups; rather, their statements about other minorities become difficult to classify, either through lack of information, or because no real attitude is present.

The low scorers, on the other hand, remain relatively consistent in the proportion of the group making prejudiced and tolerant statements—for each minority, about one-fifth of the low scorers make prejudiced statements, about half make tolerant ones. Apparently low scorers do not express tolerance as freely as high scorers express prejudice.

It is also noteworthy that the middle scorers, who probably best represent the population as a whole, are remarkably like the high scorers in their admission of prejudice although they include more individuals who make tolerant statements. This gives some justification to the attention paid to groups scoring at the extremes of the attitude scale: high scorers, at least, are not as deviant a group as might have been expected.

G. CONSISTENCY AND GENERALITY OF PREJUDICE

To apply the considerations about the validity of the interview ratings made in the section on Rating and Coding to the specific minority groups, we may begin with the relationships between the interview ratings made on the attitudes toward the various minorities and the corresponding specific scores on the ethnocentrism test.

The total score on the ethnocentrism scale was broken down into sub-scores measuring prejudice against each of the minorities discussed in the test (Negroes, Mexicans, Japanese, Jews, and general). Interview ratings were made on each of these groups in all cases where there was sufficient information to warrant a rating, using the previously discussed 7-point scale.

Correlations with corresponding sub-scores on the attitude scale, based on N 's between 27 and 64, range from .55 to .65; they are all significant at the .01 level or better, indicating satisfactory *consistency* of results from scale to interview.

Some writers have maintained that ethnic prejudices are specific responses in that attitude toward one minority group is not necessarily correlated with attitude toward another minority group. However, we have evidence of considerable *generality* of prejudice, based on both test and interview. Inter-correlations of ratings on prejudice against the various minorities made on the basis of the interview range from .69 to .97. The values obtained are lowest for correlations between attitude toward Jews and attitudes toward other minorities (.69 to .83); further substantiating the hypothesis that Jews occupy a special position as a minority group.

Similar results were obtained by inter-correlating the sub-scores on the ethnocentrism scale. Here the measure of prejudice is not subject to a rater's interpretation, and the number of subjects on which the correlations are based ($N = 600$) is much larger. These correlations range from .54 to .74. That they are somewhat lower than those obtained from interview ratings may be attributed in part to the nature of the interview rating-scale, and in part, perhaps, to halo-effects in making these ratings.

H. OPENNESS OF THE ADMISSION OF PREJUDICE

In considering attitudes towards minority groups we must differentiate the content of these attitudes (*what* is expressed) from the manner of expression (the *way* in which prejudice or its absence is manifested). It is the latter aspect with which the present section is concerned. Prejudiced or tolerant statements vary in what may be called the "openness" with which they are made, the amount of justification considered as necessary accompaniment to assertions made about minority groups.

Taking each minority separately, the child's statements as a whole were considered with respect to this definition of openness. The categorization was often made with great difficulty—a child's statements about a particular outgroup are frequently inconsistent, embodying elements of both a prejudiced and a tolerant ideology. Where it was not possible to distinguish the predominant one, or where a clear-cut attitude toward a minority group did not seem to exist, no rating was made.

Although there seems to be a partial overlap of the present classification on openness with those on "depth" of prejudice mentioned above, there is sufficient difference to warrant separate treatment.

Prejudiced responses were divided into three degrees of openness. In the first, prejudice is frankly faced, the child making no favorable statements about the minority group in question, no attempt to question the validity of any assertions made, and no qualification of unfavorable generalizations.

In the second, intermediate case, attempts are made to avoid the issue by denying knowledge of, or information about, the minority group while probing makes the presence of prejudiced attitudes apparent.

In the third kind of prejudiced statement prejudice is admitted but this is accompanied by attempts to justify the position taken. There is no real qualification of negative statements, nor an attempt to understand the minority group. The self, rather than the minority is defended. Such general statements as, "Some of them are really all right," are often made, but, coupled as they are with an affirmation of the superiority of one's own group, they represent only superficial lip service to tolerance.

We now turn to the results. Classifications of the openness of an interviewee's prejudice relative to each minority separately were expressed as percentages of the total number of those prejudiced against the particular minority.

Considering the pressures against prejudice operative in the school situation, it is not surprising that the third and last of the alternatives listed above, briefly to be characterized as "prejudice admitted with self-justification," occurs in more than half of those prejudiced against each of the five minorities.

The middle case, "initial evasion," yields no more than 10 per cent for any of the minorities.

For the first case, "frank facing of one's prejudice," the differences among the five groups are of particular interest. Prejudice is openly faced most often against the Mexicans (38 per cent of the anti-Mexicans) and the Jews (33 per cent of the anti-Semites), less often against the Negroes (22 per cent of the anti-Negroes) and the Japanese (17 per cent of the anti-Japanese), and not at all against the Chinese. Possible reasons for these differences will be discussed in the sections on specific racial stereotypes which follow.

Those expressing tolerance toward a given minority were likewise divided into three groups. Since defensiveness is not called for in the same manner as in the case of prejudice, the nature of these classifications is not exactly parallel to those just discussed.

In the first, the child admits or points to differences between the minority and his own group, but hastens to deny that these are important, or a con-

sequence of inherited disposition, that is, they are not spoken of in the context of superiority-inferiority. Sometimes the differences are even valued as something contributing desirable variety.

In the second, differences between the outgroup and the ingroup are not discussed (this may be an artifact of the particular questions asked), but the basic accepting attitude is clear from the discussion of such topics as segregation.

The third kind of tolerant attitude denies that any differences exist, insisting that the minority is "just like us." Such statements, taken at their face value, appear as a genuine expression of tolerance, but they may also be made by a child who, though basically prejudiced, has a strong need to conform, or is afraid of making aggressive statements. An error may thus be introduced by considering all such statements as tolerant. It is possible, however, for errors to have occurred in the other direction, that is, the statements of some children may have wrongly been called prejudiced.

Here again the differences among attitudes toward the five minority groups are of interest. The first of the alternatives, "differences mentioned," is most frequent relative to the Negroes (79 per cent of those tolerant toward the Negro) and least frequent about the Chinese and the Jews (about 35 per cent of those tolerant toward those groups). Corresponding figures for the middle alternative, "differences ignored," are smallest for the Negro (8 per cent) and largest for the Chinese and Japanese (44 per cent each). The third category, "differences denied," is most frequent in those friendly to the Jews (38 per cent). Some of the reasons for this will become apparent in the sequence of sections dealing with the specific content of racial stereotypes to which we turn next.

Reliability, in the sense of agreement between the two raters, is 78 per cent for the three degrees of openness of prejudice; the corresponding figure for openness of tolerance is 83 per cent.

I. THE STEREOTYPE OF THE NEGRO

If all the negative characteristics ascribed to Negroes by the children—particularly by the high and middle scorers on the ethnocentrism scale—were put together, the composite might be described something like this: a physically powerful individual, whose most pronounced features are his dark skin and large lips; addicted to drink and drugs, he is personally aggressive, attacking with razors and knives; he is destructive of property, and spends what money he has on conspicuous display of clothes and cars; he is lazy, but a good worker in menial, unpleasant jobs; jealous of whites, and

anxious for revenge on them, he provokes incidents, and threatens them by trying to "move in" and "take over," his object being to overpower and deprive them. While such a composite stereotype may actually exist in the minds of some children, the specific statements made by any one child about the Negro emphasize now one aspect, now another.

An objection may be raised about our use of the word stereotype—a word implying some distortion of reality and often used with the connotation of a myth. Are stereotypes not really deserved descriptions, and does not the average individual, having contact with members of minority groups, of his own accord come to the conclusions implied by these descriptions?

The validity of the objection is not denied outright: minority group members, because of the many pressures under which they must live, as a group may well show certain characteristic differences in behavior from the average middle-class Caucasian. Descriptions of outgroups, however, in terms not only implying the inborn nature of these differences, but exaggerating them, so as to characterize all members of a group in terms appropriate to some, or to none of those members, are properly called stereotypes, in the full sense of that word. It may require some sophistication to relate these differences to the way of life imposed on minorities by the majority, as well as to the psychological effects of discrimination, segregation, and ostracism, but it does not require more experience than that available to the average youngster to realize the error of judging individuals by fixed, inflexible standards.

Low scorers on the whole subscribe to similar stereotypes, though less frequently than do the other two groups of children. They often rise to the defense of the Negro by pointing to causative social and economic factors. Some of the children, mostly low scorers, describe Negroes merely as "nice people," or speak of them more positively as friendly and talented.

Detailed percentages as shown in Table 4 are based, within each of the three groups of subjects separately, on the number of subjects from whom specified descriptions were available (for Negroes their total is 66). None of the differences between highs and lows reaches statistical significance. (As to the reliability of this specific type of coding: Parallel coding of 19 subjects by two raters yielded 88 per cent agreement.)

One of the most striking differences between high and low scorers was not coded, but appeared with a more detailed study of some of the quotations. It concerns the child's conception of the relation between the minority group and the majority. The low scorer sees the minority as submerged and discriminated against by the majority group:

TABLE 4
PERCENTAGE OF HIGH, MIDDLE AND LOW SCORERS ASCRIBING CERTAIN CHARACTERISTICS
TO NEGROES

Characteristics	High scorers (%)	Middle scorers (%)	Low scorers (%)
Dirty	31	29	17
Lazy	8	—	4
Immoral	12	18	13
Law-breaking	4	12	4
Material success	8	12	—
Untrustworthy	4	6	—
Aggressiveness against majority	31	6	22
Aggressiveness in general	46	41	30
Different (without condemnation)	12	35	13
Clean	—	6	4
Hard-working	19	18	13
Talented	8	6	13
Other favorable descriptions	—	6	9

B34, Age 13, Low.—The reason the Negroes and Mexicans have gangs is probably because the whites don't treat them right.

B37, Age 15, Low.—There aren't many here. They are friendly, nice people. They try to be friends with people, but some white people don't want to be friends with them—don't like them. Maybe because his skin is black. But I don't see that that is a reason for that.

G32, Age 14, Low.—(Why do some people not like Negroes?) Because some Negroes don't keep their places clean and they let their kids run around dirty. Sometimes it isn't the Negro's fault because they can't get a nice place to live because the whites won't let them live there. Some Negroes are nicer than some white people.

G20, Age 14, Low.—Since there are more of them (white people) they suppress the others. I am sure that there are many Negroes and Chinese and so forth that could be doing something to help the country, but the white people won't give them a chance.

The high scorer, on the other hand, perceives the minority group as a dangerous threat to his own position, status and possessions and even projects his prejudice and hostility onto them. High scorers fear retaliation by the minority, and accuse them of the feeling of superiority the prejudiced white himself feels toward the Negro:

B13, Age 13, High.—They always seem to be getting in fights with white people, and that makes them more hated.

B18, Age 12, High.—(Segregation?) It's good in one way, bad in another. If they have to stay in one part of town they get awful mad at someone—perhaps the president—because they have to stay in one part of town and they get in fights. If they don't like another color, they wouldn't have to live with them.

G12, Age 11, High.—Because if more Negroes associated with the whites there might be race riots . . . some are very tough; some are all right. They are trying to take over the whole city. They are getting ahead of the white person. (?) Well, better jobs and better neighborhoods. They are trying to take everything away here in (name of city). . . . I hate the Negroes most because when they go to the park the colored try to take the swings away and boss you around. And the colored director there always stands up for the colored.

G29, Age 11, High.—Most of them are so cocky now. Before they had a chance to spread over the city they knew their place. Not necessarily scrubbing floors, but they didn't think they were better than other people and were not so cocky. Now they think they should rule all the city.

G15, Age 14, High.—If a Negro was president, they would probably have the upper hand. A lot of people don't treat them right. If a Negro was president, they would probably take over—get even.

The concept of being taken over and over-powered that appears in the last three quotations is more frequent among the girls (11 per cent boys; 32 per cent girls), particularly the high-scoring girls. One is reminded of the accusations of rape so often levied against the Negro male.

It is not surprising that general aggression is the characteristic most frequently mentioned about the Negro. It has often been pointed out that the lower socio-economic groups, to which Negroes mostly belong, are actually more open in the expression of hostility than members of the middle and upper classes. The three boys quoted below express this in exaggerated form.

B10, Age 11, High.—There are some good and some bad and a lot start fights. I'd like to have a couple for good fighters. They are good fighters when they fight with a knife. (?) Like somebody starts a fight and you have a gang with some Negroes to fight with you on your side with knives and guns.

B16, Age 13, Middle.—A bad thing about (name of city) . . . you have to come through where so many Negroes live. There really are some cut-throats there. Once my uncle saw a Negro that had just finished cutting another Negro's throat. He had just eaten some drug—marijuana, I think.

B25, Age 15, Low.—Most people are afraid of them for some reason. Well, they are pretty tough characters. I ran into a couple that tried to beat up on me. I used the biggest club I had near. They are tough people. I never ran across one that was weak. They can't stick up for themselves, don't know nothing about fighting. Guess it's just natural, probably because of the country they grew up in—the South.

Note in the first and third quotations the covert admiration of the physical strength of the Negro. The only difference among the three is that the low scorer appears more uncertain of his statements, contradicts himself, and makes some attempt at a causal explanation.

In the following interviews references are made to wildness, dirtiness, and laziness:

B28, Age 11, Middle.—(Segregation?) We wouldn't have to send all the Negroes, mostly those that linger around bars and a few good ones to tame them down. Don't allow any liquor bars in Africa.

B4, Age 14, High.—Oh, they're black and they have big lips and big eyes and they have an accent. I guess that's about all. (Why are they disliked?) Oh, I guess because they're so dirty all the time. And if you give them a nice modern place to live, they wreck it.

In the following protocols the Negro is considered fit only for menial labor and "dirty jobs":

B42, Age 10, High.—Some are no good, some are good for housekeeping. Well, they are black and they have wide lips.

G31, Age 13, High.—Negroes are good at working; that is, some of them. And then when they get their money, they just go and spend it on anything.

G16, Age 13, Middle.—Some races won't do jobs that others will. (?) Well, dirty jobs. The Negro and the Mexican do these. (?) They know that other people won't do it so they do it and they get a lot of pay for it.

Low scorers, in the few cases where comparable answers are given, express reluctance in making negative statements:

B29, Age 11, Low.—I don't like to say this, but a few Japanese and colored people are pretty dirty. They holler at midnight and they disturb you.

B43, Age 13, Low.—Maybe it is not right, but I think the Negroes are a little stupider than the others—because they lived so long in Africa and haven't been here long. It's not their fault.

B12, Age 10, Low.—I say they can do all kinds of work the best, such as ironing the clothes best, cleaning the house fastest, and cutting the crops down South the fastest.

Because of his alleged inferiority, the Negro is expected to "keep his place" and is criticized for any ostentatious display—in the quotations below, for having expensive cars and wearing flashy clothes:

B22, Age 11, High.—The Negroes are always trying to be smart. They have big Buicks and mostly in robberies you see about colored people, like the one in (name of city) where two colored people shot some policemen and one policeman died, I think. It just goes to show that crime doesn't pay. A policeman has to be pretty smart and quick.

G31, Age 13, High.—They're dark and they always like to start fights. I don't care much for Negroes. Some are good and some can be bad . . .

well, because they are always carrying knives around. And, well, they're dark and they put all bright colors on them.

Some children, of course, make tolerant, friendly statements. The few made by high scorers, however, generally lack the genuine, whole-hearted quality found in the records of many low scorers:

G14, Age 13, High.—They're self-reliant, and I think they're very nice. Some of them are even better than the white people.

B38, Age 14, High.—Negroes are nice people, different, I guess—they have their own ways of doing things.

B31, Age 12, High.—(Segregation?) It's bad for the state. The Negroes have taken over a lot of places. They're not as clean as we are. They have a grudge against us, but I think the grudge is our own fault, because we have taken such an awful hate against them. They don't want to go back, but they could make the South the same as this, if they worked hard enough. They should go back.

In the last quotation, note the admission of fault on the part of the whites in the plight of the Negro, with the simultaneous projection of responsibility onto the Negro.

Contrast these with the following statements made by low scorers. There is an attempt to minimize existing differences and to point out positive qualities—strength, for example, is seen, not in relation to fighting, but as an aid in superior athletic performance:

B36, Age 14, Low.—I've noticed they're very good in athletics. They have strong bodies. So far, they make very good singers. Other than that, they're the same as any other American, that is, those in California I mean. In the South many of them have grown up in different conditions and that makes them different.

B32, Age 13, Low.—They are as good as any. The only difference is color. That is no difference.

G18, Age 13, Low.—Well, they're about the same. Their complexion is dark, too. Their hair is coarse. I think that we could get along better if we lived closer together instead of separating the children, and letting the Negroes take jobs like the white man's both in the same pay.

G17, Age 13, Low.—They seem very nice to me. I haven't much acquaintance with them. Most of them seem to live in the slums. I don't know why. (Why don't people like them?) Mostly because they're black . . . I think that we could be much more tolerant and that children should be tolerant even if their parents aren't.

Note also the stress on the responsibility of the majority group for existing tensions. The last girl quoted even suggests tolerance as the particular mission of children.

J. THE STEREOTYPE OF THE MEXICAN

In the minds of prejudiced people, and perhaps also of tolerant ones, the minority group associated most closely with the Negro is the Mexican. There are many reasons for this, but certainly an important factor is skin color; some children actually refer to them as "colored people." Since the term "dirty" is applied more often to Negroes and Mexicans than to any other minority, an association between dark and dirty—sometimes observed in young children with no previous contact with dark-skinned peoples—may be inferred, as well as an association between these and what is bad or evil. The latter is understandable in the framework of our cleanliness-conscious culture; witness such sayings as "cleanliness is next to Godliness."

There are, however, differences between attitudes toward Negroes and attitudes toward Mexicans. Comparing Tables 4 and 5—reliability on the

TABLE 5
PERCENTAGE OF HIGH, MIDDLE AND LOW SCORERS ASCRIBING CERTAIN CHARACTERISTICS TO MEXICANS

Characteristics	High scorers (%)	Middle scorers (%)	Low scorers (%)
Dirty	18	8	12
Lazy	6	25	6
Immoral	18	8	12
Law-breaking	12	8	12
Material success	—	8	—
Untrustworthy	6	—	—
Aggressive against majority	6	—	6
Aggressive in general	24	8	25
Different (without condemnation)	24	25	25
Clean	6	—	6
Hard-working	18	17	12
Talented	12	8	6
Other favorable descriptions	—	8	12

latter is 86 per cent for 12 subjects coded by two raters—we note that in the description of the Mexican there is less reference to aggressive behavior than in that of the Negro, and almost none to aggressiveness directed specifically against the majority; that is, the Mexican is not perceived as a national threat. But he is sometimes described as dirty, immoral (e.g., drinking), and lazy, the latter two characteristics often in the context of pleasure-seeking.

This difference in stereotype may be partly due to differences in the position of these two groups in our national life. To detail these differences is beyond the scope of this paper. We would like to point, however, to

such factors as numerical strength (approximately fifteen million Negroes, three million Mexicans), the length of time during which the two groups have been important as minorities (the Mexican is a relatively recent addition), the concentration of Mexicans in certain areas of the country, and skin-color which, though a bond of similarity, is also a differentiating characteristic in many cases. The area in which our subjects live has a relatively large Mexican population; attitudes toward this group are thus likely to be more crystallized than attitudes of children in some other parts of the country.

The following interview excerpts describe some of the characteristics discussed above:

B24, Age 11, High.—They're the same color as the Negroes. They talk different from us and they have different kinds of houses. They have a different style . . . they don't dress the same way as we do. They play fast music.

B12, Age 10, Low.—Well, they're colored people mostly, and some are very cruel and mean. When they take a club and hit you, you don't know what is happening because they do it so hard.

B28, Age 11, Middle.—I wouldn't want a Negro or Mexican (for doctor) because they seem to have dirty hands all the time.

G19, Age 13, High.—A lot of them drink a lot. I haven't known any, I've just heard this. They're not too social. (?) They don't go around, they kind of stay among themselves.

B21, Age 13, Middle.—They are lazy. Some aren't but most of them are. They just sit around and have fiestas, but they're not dirty.

The last quotation seems to contain an implicit envy of the minority group, which is able to enjoy pleasures denied to the majority.

The first of the following quotations from the interviews of low scorers gives a good example of negative characteristics ascribed to a group in such a manner that a tolerant answer results; the boy stresses the relationships that may exist between the kind of behavior described and unfavorable situational factors. In fact, all three of the quotations may remind one of the social scientist's approach to the problem, a feature previously noted in low scorers. The last girl in particular sees the situation with a clarity and penetration unusual for children. Note especially her willingness to tolerate differences of which she personally disapproves.

B6, Age 14, Low.—Some of them are rowdies. They aren't nice. Some Mexicans are picked on even if they don't deserve it. So then they do things to give people a reason to be mad. Probably they got a bad reputation because of a few bad ones that came up from Mexico.

B14, Age 14, Low.—I have read of Mexicans knifing people. It seems

to me that perhaps their ancestors used knives—in Mexico a lot of them carry knives, can't afford a gun. I think they are what we call lower class, as far as wealth goes. If you went to Mexico among the peons, you would find that they use knives mostly for cutting weeds and things like that. Some Mexicans who came here might figure they needed some defense and rely on knives.

G20, Age 14, Low.—The Mexicans haven't been treated decently, so they have acted in a bad way. I think it wasn't their fault. If we would give them equal chances, they wouldn't be in any more trouble than anybody else. The Mexicans came here thinking they would have better chances to get along, but weren't treated as equal of the whites. They think, "If the Americans don't like me, I might as well deserve it". Here there are a lot (of young people) that smoke and do things that some people don't like. One set here—Pachukes—are in a different class and have their fun in a different way. I haven't anything against them, but I wouldn't go around with them because I wouldn't want to do the things they do.

K. THE STEREOTYPES OF THE CHINESE AND JAPANESE

Attitudes toward these two groups in the 1946-47 period were probably influenced by the events of the war, toward one as aggressor, the other as victim of aggression. This is particularly true for West-coast children who, living among relatively large concentrations of these minorities, have seen the mass evacuation of a large number of Japanese or of citizens with Japanese ancestry.

Unfortunately the interview material obtained on attitudes toward these two groups was extremely sketchy—less than half the subjects were questioned about them, and of these, few gave answers clear enough to make inferences about the direction—or even presence—of an attitude. Descriptions of Chinese and Japanese were yielded by only about 30 subjects each. Tables are therefore not shown; the material below is a qualitative summary of statements made about the two groups.

Of all minority groups—except, in some cases, the Jews—our children feel the closest kinship to the Chinese. This may result from the apparent adoption by the Chinese of the more obvious middle-class values—cleanliness, industriousness, and thrift. Thus the Chinese are described as clean (13 per cent), hard-working (17 per cent), and—by implication—quiet, patient, and submissive. Only one child makes reference to aggressive behavior of any kind.

B21, Age 13, Middle.—They are hard workers. They jump at a chance to get an education. They are poor people, too.

B29, Age 11, Low.—The Chinese are probably the cleanest and the

neatest and they think of other people more. They wouldn't holler at night and such things.

B38, Age 14, High.—Their language and books are different. They don't live in the same kind of houses as we do. (?) It's all right to live in that kind of house if they want to.

B37, Age 15, Low.—All I've known are nice people. They still use a lot of Chinese ways, but the young ones are more American than anything else. Lots of times their parents don't speak anything but Chinese.

As recognized by the last boy quoted above, two different sub-cultures co-exist among the Chinese in the United States. The next quotation shows some feelings of strangeness and mystery ascribed to the old culture; this is used by the following girl in the service of a justification for prejudice.

G16, Age 13, Middle.—They seem sort of quiet. They are about the same as we are, only they dress differently. . . . Some Chinese have crazy ideas. (?) Well, I've heard this, maybe it isn't true. That a lot of people go to Chinatown and are found missing. Once when my mother and I were in a car going through Chinatown we were nearly run over by a Chinese in a car. I don't think they like Americans very much.

As far as the Japanese are concerned, the most frequent accusations are those of aggressiveness (19 per cent) or of treacherousness and sneaky behavior (29 per cent). These characterizations existed before the war but were certainly intensified by it. The stereotype is well illustrated by the distinction, made by the high-scoring girl below, between "Japs" and "Japanese."

G38, Age 10, High.—There are a lot of Japs and Japanese who are butchers. (?) I guess they have different blood. They almost look alike. Their eyes are alike. There are not too many Japs, but there are a lot of Japanese.

G15, Age 14, High.—The Japanese seem sort of sneaky, because of what they did in the war.

Different in quality is the answer of this low-scoring girl who specifically fights pre-judgment and makes an issue of the necessity of obtaining facts before forming opinions.

G17, Age 13, Low.—(Why don't some people like them?) Because they collaborated in the making of the war, and because they're yellow and have different beliefs. It's funny that so many children say, "the dirty Japs" and they don't even go over there and see them. So many people here hate them when they don't even know what it's all about.

Another objection is offered by the next high-scoring boy, that of am-

biguity. The fear of not being able to recognize members of minority groups appears again in the discussion of Jews.

B22, Age 11, High.—(Who would you not want as a neighbor?) Japanese. I couldn't tell if they were Japs or Chinese.

L. THE STEREOTYPE OF THE JEW

A survey of the stereotype of the Jew as conceived by our subjects is given in Table 6. In accordance with the difference in stereotype, the categories are somewhat different from those used for Negroes and Mexicans. As to reliability of coding, the two raters showed an agreement of 75 per cent for 15 cases compared. Total of subjects for Table 6 is 43.

TABLE 6
PERCENTAGE OF HIGH, MIDDLE AND LOW SCORERS ASCRIBING CERTAIN CHARACTERISTICS TO JEWS

Characteristics	High scorers (%)	Middle scorers (%)	Low scorers (%)
Dirty	—	7	5
Cheating in money	38	33	23
Material success	38	13	14
Untrustworthy	6	13	—
Aggressive against majority	25	—	—
Aggressive in general	6	7	—
Different (without condemnation)	25	33	45
Talented	12	7	—
Other favorable descriptions	6	—	5

The most frequently mentioned of the characteristics ascribed to the Jew is economic success and cheating in money matters. More specifically, cheating is mentioned by 38 per cent of the high scorers and 23 per cent of the low scorers, while material success is brought up by 38 per cent of the high scorers and 14 per cent of the low scorers. This finding is of special interest, since other data show that material success is a crucial goal for the prejudiced person. A few of the children—high-scorers only—also accuse the Jew of aggressiveness against the majority group. Mention of difference in a neutral context, i.e., without condemnation, is more frequent among those scoring low on the attitude scale (25 per cent high scorers, 45 per cent low scorers); this category contains references to such things as customs and religion.

It is interesting to note that most of the characteristics ascribed to Jews are, in contrast with those attributed to Negroes and Mexicans, not in open conflict with middle-class values. There seems to be a double standard

operative here—the success ethic, so admired in our culture, is the very thing condemned in the Jew. This stereotype is probably reinforced by the obvious fact that the Jew is more often allowed to succeed in competition with the non-Jew than are members of any other minority group.

In the following quotations the Jew is attacked or pointed out for his success, his social and economic status, his control of money and of the sources of pleasure (theaters, movies). In the third quotation a low-scoring boy, though ambivalent in his response, expresses the stereotype clearly.

B4, Age 14, High.—Oh, they say that they own the banks and everything. And they darn near own the country. (What do they do best?) Making money. They're good at owning the movies.

G17, Age 13, Low.—Oh, people sometimes call others a Jew with a derogatory tone. Some boys will say, "Oh, you're an old Jew." It has something to do with money.

B12, Age 10, Low.—Well, they're poor the same as I am. Sometimes there are Jews that are rich, the richest people in the world in a manner of speaking.

G16, Age 13, Middle.—All are about the same. A lot of Jews are presidents and bosses of business. (?) Well, they seem to be able to get ahead faster (said pleasantly). There are good and bad in every people. I imagine that Jews are like that. Some are good and some are bad. There are more actors and actresses among the Jewish people. The Jews seem to get ahead and have more talent.

The last quotation also contains references to the expressiveness of the Jews, ascribing special dramatic talent to them. The following two girls bring up the negative side of expressiveness—proneness to quick anger, change of mood, and ready exhibition of feelings. The emphasis on "control" in our culture in general, often exaggerated in the prejudiced home, is likely to induce both contempt and envy of those who permit themselves greater expressiveness.

G29, Age 11, High.—They all change too quickly. One minute they like you and the next minute there's a big argument. I wouldn't want to get too friendly with one. There's a Jewish girl in my class that is nice and all that, but I wouldn't want her for my best friend. (?) A lot of people say they are too cheap, they try to make one dollar and then they look for another dollar underneath.

G32, Age 15, Low.—I know only one. She doesn't get along with her husband. They seem to get mad at every little thing. When they get married I don't think they want to stay married long. They like to have their own way in every little thing. . . . I don't see much difference in them and I think they ought to live everywhere. They seem just as neat and clean as other races.

The next two high scorers go on to criticize the Jew for exhibitionism, showing off of possessions and wealth. They refer to domineering qualities (implied in the discussion of business success), the boy even to the extent of accusing them of making slaves of the Gentiles.

G31, Age 13, High.—Oh, they say they're selfish and they tell you what to do. They have a lot of clothes and like to be the big one around.

B10, Age 11, High.—They think they are smart and go anywhere they please. They think they're hot! They dress up in all kinds of jewelry. Some just kidnap girls and boys and use them for slaves.

The Jew is sometimes accused of a manipulation of others to his own financial advantage. In the highly emotional response below Jews are seen as capable of taking subtle, terrible revenge.

B16, Age 13, Middle.—I know a Jewish boy. He is real nice. Some Jews shouldn't even be called a Jew because they are awful nice. (?) Usually people call them tight wads. They try to get everything they can. Once I saw a radio program. The man that owned the station was a Jew. A regular mob of people came and tried to get a free plug on the radio. The radio man said, "You Jews are all alike." When one Jew jumps another Jew, they get scared. I think a Negro would cut another person's throat. A Jew won't even have to do that. He will lead them to a terrible death by taking everything they've got in some awful way like gambling.

Less specifically, the Jew is feared because he is somehow "different," though there is no clear conception of the nature of the difference.

B4, Age 14, High.—Well, they're different people. I don't think they're a different religion, but they are a different race.

B13, Age 13, High.—They're not any good. Well, my pop said they're not any good. I just took it for granted I didn't like them.

G38, Age 10, High.—Well, they are different. (?) They look different than other people. Most of them are selfish. A lot of them cheat a lot.

B12, Age 10, Low.—Well, they are white people and I wouldn't like to have them for friends. They can go to any church they want to, and they are just as good as we are. But I still don't like them. Well, they speak such a language that nobody can understand, and they don't want to learn the American language . . . well, because of their language and because of the clothes they wear. They wear mostly black which we don't do.

Not only must Jews, relative to other minorities, seem successful in the economic world; at the same time, they are lacking in distinguishing characteristics so that they are not always recognizable. This latter fact seems to lead in some cases to a reduction of the general level of prejudice,

in others, to its magnification. The difficulty in distinguishing Jews from Gentiles constitutes an anxiety-producing situation, especially to some prejudiced people with their proneness to make black-white discriminations and to shun any kind of ambiguity.

There is, moreover, fear of mobility; if the Jew is not isolated and restricted in his ability to move, his detection becomes even more difficult. Thus one of the high-scoring boys criticizes, and perhaps envies, Jews because they "go anywhere they please." This particular boy is a fearful child who seldom dares go alone more than a few blocks from home.

Other children stress differences in belief; the first boy quoted below expresses also the fear of ambiguity mentioned above.

B28, Age 11, Middle.—They kind of believe in their own religion. They are no different than we are, but it would be better if they went back to their own country. You can't very well tell them from Americans. There are a lot in our neighborhood. Sometimes we are liable to say "You are a Jew" to an American and it insults him because the word Jew has a bad reputation. (?) There are bad Jews and they give a bad reputation to Jews so you think all Jews are bad.

B26, Age 11, High.—They believe in different things; they have different New Years. They don't believe in Christmas.

That the absence of distinguishing characteristics can also be the basis for unprejudiced statements is shown by the quotations below:

B43, Age 13, Low.—You can't tell them from others . . . a lot of people think the Jews are stingy, but I don't know. Well, a lot of white people are stingy too—a lot are no matter what creed.

B30, Age 10, Low.—It's silly to say that. If you cheated somebody, maybe a Negro, it's the same as if a Negro boy calls another Negro a white boy. We're all the same. My father has known Jewish people that are more honest than other people that he has worked with. I know some Negroes, and as far as a Jew, I can't tell whether a person is a Jew or not.

Because the Jew is not clearly set apart from the Gentile, there is often confusion about the meaning of the word and the identity of the group. Our children seem to have failed to develop crystallized opinions about Jews more often than about any other minority discussed. Probing in many cases reveals only the absence of a definite attitude.

One occasional stereotype of the Jew is not found in our subjects, that of the Jew as a radical and agitator. The Jew, both because of his social position and because of the ambiguity of his presence, lends himself particularly well to projection. Thus, to our predominantly lower-middle-class sub-

jects, the Jew appears as the unduly rich; by wealthier subjects, or by those aspiring to wealth, he might be perceived as the challenger of the economically privileged. While some of the middle-class reject those below themselves, e.g., Negroes, primarily because of the threat they are to their precarious status, they may reject those above for having the things they themselves are striving, all too often unsuccessfully, to attain.

M. SEGREGATION OF MINORITIES AND ITS RATIONALIZATION

It is relatively easy to pay lip-service to tolerant beliefs, but it is more difficult to assert that one will behave tolerantly in a specific situation. Thus a discussion of segregation practices in the interviews proved more evocative of verbally expressed prejudice than any of the other approaches used in this study.

Of a total of 16 high scorers on the ethnocentrism scale from whom we have material on segregation, three-fourths favor exclusion of at least some minority groups from certain residential areas; among 22 low scorers there are only three who do likewise. A few of the segregationists favor added exclusion from parts of the country, and about $1/3$ of them from the country entirely. The difference between high and low scorers is significant at the .001 level.

Only one of the 16 high scorers but 14 of the 22 low scorers concerned openly and unqualifiedly disapprove of segregation practices. The difference is significant at the .01 level. A few more would condone segregation only if specifically chosen by the minority in question.

It is of interest to point out here, as was done in the section on openness of admission of prejudice, the similarity of the middle scorers to the high scorers. Of the small number questioned ($N = 11$) nine favor segregation and none show unqualified disapproval.

For segregation of Negroes specifically, the totals questioned are 12 and 13 respectively in each of the three groups. Nearly all high scorers, about two-thirds of the middle scorers, and less than one-half of the low scorers favor segregation of some sort. The rest are qualified or unqualified anti-segregationists.

The children's actual statements will now be considered. A few children are able to express their segregation beliefs openly and honestly. Those quoted below show an unusual degree of blandness:

G12, Age 11, High.—We prefer our own race first, and I prefer the the Greek next. I hate most colored people. After the Negroes I hate the Mexicans and Spanish.

G14, Age 13, High.—I would resent living next to somebody that was a foreigner . . . they should go back where they came from. (?) To the country they came from. Another thing too, some of them were born in America, a lot of them. I still think they should go back where they came from.

B21, Age 13, Middle.—The Negroes should go back to where they belong. To Africa. The Americans should just stay right here.

B12, Age 10, Low.—Well, the people in the country should just make a law to chase them (Jews) out when they come in.

Glorification of the in-group is also pronounced in these quotations. In the following quotations it takes the surprising position of making displaced persons responsible for the war, absolving the Germans of responsibility. In one case, the third quoted below, there is agreement with Hitler's actions toward the Jews. This boy is a good example of a child who obtained a low score on the ethnocentrism scale, not because his general level of prejudice is low, but because, while showing an intense aversion toward one minority group, he is relatively indifferent to the others (see also *B12* above).

B24, Age 11, High.—(Segregation of displaced persons?) Well, all the others have their countries. Well, they have their style in their country, and we have our way of doing things . . . it was probably their fault that the war started. And they got what was coming to them when their houses were bombed.

B31, Age 12, High.—I think everyone should stay where he should be, where he was born. Now, the Dutch haven't hurt us a bit. The Irish haven't done us any harm at all. The Germans, just in California, haven't hurt us. My grandmother was German. She's dead now. They're the closest you could come to an American. They didn't start the war.

B35, Age 13, Low.—(Jews?) Hitler, Caesar and Napoleon had the right idea. They were all trying to get rid of them. Caesar fed them to the lions; Napoleon put them in front of the firing squads; Hitler put them in front of the firing squads and also burned them alive. . . . Segregation is a good idea because most people don't like Jews. It's about religion or further back in history. Some don't like them just because they don't like them.

Other children, to be quoted below, must clothe their advocacy of segregation with various attempts at justification. Among this group, four kinds of response can be distinguished: the projection of aggression, the projection of segregation wishes, the projection of responsibility, and the patriarchal justification.

The first kind is relatively open in the expression of prejudice, arguing that minorities must be segregated because they constitute a threat to the

majority group. The threat is usually seen as one of aggressive behavior. The following are examples of this kind of justification.

B10, Age 11, High.—I wouldn't mind having all the Negroes in (name of city) and all the white people in a different state. (?) The Negroes should live in one place and the whites in another so there won't be no trouble. Negroes make trouble, start a war.

B26, Age 11, High.—Oh, they could have reservations for them (Mexicans) like they do for the Indians; then they wouldn't try and start nothing.

B4, Age 14, High.—(Should people of different races live together?) No, no I don't. Because they get in fights. And if they lived together—well, they'd go to their own parties.

The second group justifies segregation by projecting the demand for the desired segregation onto the minority group in question, thus releasing the majority from responsibility for discriminatory practices. By this pretense of objectivity, tolerance, and acting for the good of the minority, these children are able to favor segregation with the approval of their conscience.

B31, Age 12, High.—I don't think they (Japanese) are bad. They are something like the Chinese. I think they are clean, but they should be in their place. They shouldn't mingle in our crowd, and go to social gatherings where we go. I don't think we should go to theirs, either. I think they feel better if all their race is together, the same as we would feel.

G15, Age 14, High.—I think it should be that way. Well, if I have a home, I wouldn't particularly want to have a Negro neighbor, not that I think they are not as good as we are. It is for their good as well as ours. They would probably like better to live where all are Negroes, or say Japanese. I think it's a good idea. Meeting them is a different thing. I suppose I shouldn't say that.

G16, Age 13, Middle.—I think it is better that way, except for the Jews. Everybody is a little different and gets along better by living with their own people. The Negroes, especially, should be segregated, and the Chinese to some extent.

G2, Age 13, Middle.—It's better for them and better for us.

G18, Age 13, Low.—Well, in some places. Well, where there are quite a few white people and not so many Negroes. Schools that have Negroes, they fight the white children and then they don't like Negroes when they grow up and the white children sort of call them names behind their backs. If they had separate schools, white children would like them better.

A sub-variety of the second type, relatively rare, projects responsibility for the demand for segregation to other members of the majority group; thus

the formula becomes: I don't object to non-segregation, but I must protect my friend who does.

B16, Age 13, Middle.—I think we should stay by ourselves. It is the best policy. Maybe your friend wouldn't like a colored person. For example, if you lived around them.

A third group of children attempts to do away with guilt feelings by giving minorities every advantage, depriving them of nothing save the right to live among members of the majority group. This is the well-known formula of separate but equal opportunities.

B28, Age 11, Middle.—We should send the Mexicans back . . . make it the same with the Japanese. Make things better with the Chinese. Let the Negroes get on a ship and take them to Africa and let them settle down and have a modern country there in Africa. For all we should make good living conditions where they go.

Some of the last-quoted answers may seem to approach a measure of tolerance. Yet they are a long way from genuinely tolerant responses as given by the children quoted below. These passages show a real respect for the rights of others, even of those against whom criticism is raised. All of the children quoted oppose segregation, but they do so for different reasons.

Some would condone segregation only in the specific instance of segregation being chosen voluntarily by the members of the minority, perhaps with an implicit recognition of the tendency of people of like backgrounds and interests to form social groups.

B30, Age 10, Low.—It is up to the people whether they live with their own group or among peoples of another group. I wouldn't mind any nationality in my neighborhood.

B32, Age 13, Low.—If they want to, it's okay. If they are forced to, it isn't right.

B43, Age 13, Low.—It's all right if the people want it, but if they want to scatter around, there should be no law against it.

Others, more naively, ridicule the policy of segregation:

G13, Age 14, Low.—I think it is sort of silly. I don't see why they want such rules. We should let everyone live where they wish. Every race should have the same privilege. They are Americans like the rest of us.

G27, Age 13, Low.—It's silly. Personally, I don't think it is right.

Some point to the illegality of segregation:

B36, Age 14, Low.—Well, I suppose there is a reason for it, but I can't tell if it is good or bad. As far as I am concerned, it is bad. There shouldn't be any reason why people should be restricted if they're American citizens.

B41, Age 12, Middle.—Well, I see a few (Negroes). They seem nice. They don't bother anybody. They have as much right to live here as the whites. The white hasn't any right to say which nation can live here.

Some children object to segregation for moral reasons. The girl quoted below turns to the Bible for support, and to the principles on which this country was built.

B34, Age 13, Low.—The colored people have such lousy homes. They need new houses in good districts.

B6, Age 14, Low.—It's wrong! I guess it's because they look different. There is no other reason. All people are of the same intelligence if they have the same amount of learning.

G36, Age 13, Low.—Oh, some say they just don't like them (Negroes) because of their color, but it says in the Bible that all men are created equal. Some say that they should be taken back to their own land, but we brought them over here and made them slaves and I think they should stay here because America is supposed to be a country where all different nationalities and religions get together.

Another girl points to economic factors involved in segregation.

G20, Age 14, Low.—Here the races live most places. But of course the Negroes can't get as good jobs so can't live in the better sections where it costs so much. (Do Negroes prefer segregation?) I don't think so. They would rather be equal to everyone else.

Several children refer to the bad psychological results of discriminatory practices, recognizing that hate may breed hate. The last girl quoted makes an eloquent plea for people to learn to live together, even if they have difficulty in doing so.

B37, Age 15, Low.—I don't believe in it. If you separate people, they don't feel as friendly to you as if they lived near you.

G32, Age 14, Low.—I don't think they should have those rules—like where we live now they don't allow the Spanish people there. I think they ought to have just as much right to buy property there as anybody. Those that live in special areas get to not like other people, get so mad because they don't know other people. They ought to be all broken up and let all races mix together.

G7, Age 14, Low.—It's good for the country to have all races mixed

together. If they can't get along together, well, I don't think they should be separated, but should learn to get along. (Segregation in cities?) No, I don't agree that it is good, because they are separated into different groups and don't learn to see things the same way and so get into arguments. It might even cause civil war again.

One of the low-scoring boys even suggests that we segregate the very people who demand segregation.

B2, Age 12, Low.—Some people think the Jews shouldn't come into America. People that think that should be eliminated. Have the people that say that go out of the country and let them come back when they don't say it any more.

N. FAMILY BACKGROUND AND PREJUDICE

As was pointed out in the introduction, interviews were also conducted with parents of the children interviewed. We did not aim at a complete coverage of the parents, however. Thus, while all the parents were approached in this matter, the ones actually interviewed were those with whom arrangements could be made for the time periods available to the interviewers. Interviews were completed with the mother or with both parents of 43 of our 81 child subjects. These interviews were rated for attitudes towards Negroes, Jews, and for general prejudice, on the same seven-point scale which was used for the children's interviews. The ratings were again done blind. The parent interviews were less comprehensive than those obtained from the children and it was not possible to make all ratings. The parent ratings were made by a rater who had not rated the children. Prejudice ratings of the interviews with the children were correlated with both their fathers' and their mothers' interview prejudice ratings. This procedure was repeated for three kinds of ratings, that is, prejudice against Negroes, prejudice against Jews, and general prejudice. The six correlation coefficients thus computed are in part based on very small numbers of comparisons, the *N*'s ranging from 9 to 41. Although not much in the way of statistical significance can be claimed under these circumstances, we nonetheless wish to report that the obtained values of the coefficients are all positive, ranging from about .2 to about .6. The overall picture is suggestive of intrafamilial influences upon ethnocentrism in the child, although there can be little doubt that other factors are also operative. As subsequent articles in the present series will show, the attitude toward minority groups is only one of many attitudes likely to be transmitted within the family. This raises the question as to whether prejudice is transmitted

directly, or else indirectly via a more general formation of character. In the latter case the child's attitudes toward minority groups would be a concomitant of shared need structure within which prejudice fulfills a definite function, rather than being acquired in a more incidental manner.

O. SUMMARY AND CONCLUSIONS

Eighty-one white American Gentile children, 10 to 15 years of age, were interviewed about their attitudes toward five minority groups: Negroes, Mexicans, Chinese, Japanese, and Jews. The subjects were selected from a larger group on the basis of a scale designed to measure ethnocentrism. Twenty-seven of the children interviewed are defined as high scorers (i.e., they are among the most prejudiced in our total group), 36 as low scorers, and 18 as middle scorers in terms of this scale.

In spite of the difference in approach in the coded interviews and the tests, there is substantial agreement between them. Prejudice is measured by the total ethnocentrism scale score correlates .67 with an over-all interview rating on general prejudice. This consistency of prejudice from one instrument to another is only one expression of its generality. Another is the high inter-correlations (.69 to .97) among ratings on prejudice toward various minority groups made on the basis of the interviews. The objection that a halo-effect was operative in making these ratings can be answered, at least in part, by reference to similar inter-correlations between corresponding sub-scores on the Ethnocentrism scale (.54 to .74).

An interview statement of a negative nature about minority groups was considered to reveal prejudice when there was no attempt at rational explanation, or when qualifications of the statements represented an attempt to justify the self rather than to understand the minority. The latter kind of statement occurred relatively frequently, probably because the schools in which the interviewing was done take a strong stand in favor of a tolerant ideology. On the other hand, statements were considered as expressions of tolerance when an effort was made to explain any negative characteristics ascribed to minority group members by reference to causal social factors, or when the inherent character of these differences was denied and their importance minimized.

Ninety-six per cent of the children scoring high on the ethnocentrism scale and 22 per cent of those scoring low made prejudiced statements about Negroes. Eighty-three per cent of the middle scorers expressed prejudice against the Negro. Similar results were obtained for the other four minorities, except that fewer of the high scorers made prejudiced statements

about these groups, and more of them made statements unclassifiable on a prejudice-tolerance dimension. Among the more frequent characterizations of the Negro, particularly by high scorers, are those describing him as hostile towards the major group, as dirty, and as immoral, but also as hard-working (Table 4). But it must be noted that no children spontaneously list all these attributes and that the low scorers subscribe, though less often, to the same stereotypes. Furthermore, not one of the characteristics listed is mentioned by more than a minority. In spite of all this, the characteristics assigned to Negroes, especially among high scorers, seem to imply the prevalence of what we may conceive as a stereotype.

To a somewhat lesser extent, terms similar to those applied to the Negro are applied also to the Mexican. Notably there is less mention of aggressiveness and particularly of aggressiveness toward the majority than in the case of the Negro (Table 5). Such factors as difference in numerical strength between Mexican and Negro and the relative importance of the two groups in our national life, may in part account for the difference in stereotype.

The stereotype of the Japanese, still less pronounced than that of the Mexican, includes the attributes "aggressive," "treacherous," and "sneaky." This stereotype was probably intensified by the war. The Chinese, on the other hand, is more often considered as clean, hard-working, and non-aggressive, traits commonly valued in a middle-class culture.

The Jew has a peculiar position relative to other minority groups. Unspecified differences between Jews and Gentiles, although prominently asserted by some of the subjects, are at the same time more often explicitly denied than is the case with other minorities; but prejudice is often expressed with unusual violence against this group. Jews are described by many subjects, especially by high scorers, as economically successful and as cheats in money matters (Table 6). There seems to be a double standard of evaluation here, material success, for example, being considered a virtue in members of the majority but a vice in others. The records of some of the high-scoring children allow us to trace resentment against the Jew to his imputed success and status, his dominant position, unrestrained mobility, expressiveness, exhibition of feelings, and of wealth. The difficulty of recognizing the Jew as compared with members of other minority groups and his relatively greater freedom of social intercourse with the majority, goes in some cases with a reduction of the general level of prejudice, in others with its magnification.

High and low scorers also differ in their conception of the position of the minority group in relation to the majority. The low scorer sees the minority

as discriminated against, submerged, deprived and frustrated by the majority group. The high scorer emphasizes the possible threat and danger of minority groups to his own and the resulting necessity of controlling them lest they take over and revenge themselves on the majority.

The significantly greater number of prejudiced statements made by high-scoring children in response to inquiries about minority groups does not necessarily imply that these problems are more on their minds, nor that they are more willing to discuss them voluntarily. On the contrary, we find that in answering the more general question, "What is an American?", it is the low-scoring child who more often mentions the issue of race, stressing at the same time that race is not a basis for excluding anyone from being an American. The high-scoring subject often seems more reluctant and embarrassed about bringing up the topic spontaneously, at least in the school setting in which the interviewing was done. Difference of birth or national background is sometimes stressed as being incompatible with being an American. A few children go so far in their requirements as to stress external characteristics such as health and cleanliness.

A tendency toward dichotomizing, exaggerated adherence to conventional values, the avoidance of insight into the characteristics of the in-group and fear of threats and dangers from the out-group were themes found in the discussion of minority groups by the high scorers. They will appear again when we discuss the children's conception of the social scene in general, and their attitudes toward interpersonal relationships. These topics will be considered in succeeding papers.

Questions pertaining to segregation, that is, to specific actions against, rather than mere attitudes toward, minority groups showed the clearest differences between high and low scorers. About three-fourths of the high scorers and middle scorers questioned, as against a scattered few of the low scorers, favor segregation. This topic elicited the most open expressions of hate on the part of some high scorers, but also brought out the most eloquent pleas for the rights of minorities on the part of some low scorers. Some of the latter children are very explicit in pointing out that the very qualities often resented in the minority group are specifically bred by such measures as segregation.

Although our material is somewhat scant in this respect, correlations between the children's and their parents' prejudice, between .2 and .6, point in the direction of family influences in the area of ethnic tolerance or intolerance.

Our results have, on the whole, some distressing implications for those interested in combating prejudice. The children's attitudes are usually a mixture of prejudiced and tolerant ideologies. Our high-scoring subjects, on the whole, were freer in expressing prejudice than the low-scoring children were in expressing tolerance. Furthermore, the middle scorers who are more representative of the bulk of the population show great similarity in their responses to the high scorers. Thus the attitudes of the children as a group are predominantly prejudiced ones.

Though the prejudiced child shows, in many ways, an exaggeration of values which are accepted in our society, it should not be overlooked that the low scorer too is a part of this same society. The implicit or explicit antagonism between these two opposing ideologies characterizes not only our culture as a whole, but, to varying degrees, every single individual. How this struggle will end does not hinge on psychological factors in isolation. As such factors are in the end manifestations of broader cultural influences, it is only by an understanding of the interplay of economic, sociological, and psychological phenomena in their entirety that a full appraisal of the relative strength of these two opposing trends as well as of their nearness to, or remoteness from, the overt behavioral level, can be made.

REFERENCES

1. ADORNO, T. W., FRENKEL-BRUNSWIK, E., LEVINSON, D. J., & SANFORD, R.N. The Authoritarian Personality. New York: Harper, 1950. Pp. 976.
2. FRENKEL-BRUNSWIK, E. Motivation and behavior. *Genet. Psychol. Monog.*, 1942, **26**, 121-265.
3. ———. Dynamic and cognitive categorization of qualitative material: I. General problems and the thematic apperception test; II. Interviews of the ethnically prejudiced. *J. of Psychol.*, 1948, **25**, 253-260, 261-277.
4. ———. A study of prejudice in children. *Human Rel.*, 1948, **1**, 295-306.
5. ———. Intolerance of ambiguity as an emotional and perceptual personality variable. (In: Symposium on Interrelationships between Perception and Personality, Part I). *J. Personal.*, 1949, **18**, 108-143.
6. HOROWITZ, E. L. The development of attitudes toward the Negro. *Arch. of Psychol.*, 1936, No. 194.
7. ———. Racial aspects of self-identification in nursery school children. *J. of Psychol.*, 1939, **7**, 10-99.
8. LEVINSON, D. J. An approach to the theory and measurement of ethnocentric ideology. *J. of Psychol.*, 1949, **28**, 19-30.
9. MINARD, R. D. Race attitudes of Iowa children. *Univ. Iowa Stud. Charac.*, 1931, **4**, No. 2. Pp. 101.
10. PETERSON, R. C. Motion pictures and the social attitudes of children. (In: Chicago Assoc. for Child Study and Parent Education. *Developing Attitudes in children.*) Chicago: Univ. Chicago Press, 1933.

11. RADKE, M., TRAGER, H., & DAVIS, H. Social perceptions and attitudes of children. *Genet. Psychol. Monog.*, 1949, **40**, 327-447.
12. ZELIGS, R., & HENDRICKSON, G. Checking the social distance technique through personal interview. *Sociol. & Soc. Res.*, 1934, **18**, 420-430.
13. ———. Racial attitudes of children. *Sociol. & Soc. Res.*, 1937, **21**, 361-371.

Institute of Child Welfare
University of California
2739 Bancroft Way
Berkeley 4, California

SHORT ARTICLES AND NOTES

The Journal of Genetic Psychology, 1953, 82, 137-142.

CHILDHOOD FEARS IN RELATION TO HYPNOTIC AGE REGRESSION: A CASE REPORT*¹

Westchester County Department of Health, White Plains, New York

MILTON V. KLINE

A. INTRODUCTION

Recent studies by True (9), Mercer and Gibson (7), Leeds (6), and Kline and Guze (4) have tended to confirm the hypothesis that hypnotic age regression is fact and not artifact as has been held by some hypnosis investigators. Apart from the problem of the validity of hypnotic age regression its rôle in hypnotherapy has generally been considered as a valuable technique in appropriate psychotherapeutic settings. Bergman, Graham, and Leavitt (1), Mercer and Gibson (7), Spiegel, Shor, and Fishman (8), and Kline and Guze (4) have presented clinical and experimental evidence of the utility of psychodynamic data obtained with hypnotic age regression pertinent to psychodiagnosis and therapy.

The suggestion has frequently been made that hypnotic age regression might be subjected to more intensive psychogenetic or developmental studies of longitudinal nature. This would provide additional data on the validity of the phenomena and might conceivably extend the use of regression techniques to experimental studies of personality growth. This paper reports on the comparative reaction to experimental fear situations by children at early ages and a subject hypnotically regressed to a similar age level. As such, this report is a first of a series dealing with the relationship of hypnotic age regression to psychogenetic aspects of emotions and emotional expression.

B. PROCEDURE

From Jersild's (3) *Studies of Children's Fears* some experimental situations were selected and a 25-year-old female subject placed in hypnotic age regression at a level of three years was exposed to the fear producing situations in a manner similar to that described by Jersild in his original work with

*Received in the Editorial Office on January 31, 1951.

¹Read before the Society for Clinical and Experimental Hypnosis, January, 1951.

children of that age. The experimental situations and Jersild's description of them are as follows:

1. **BEING LEFT ALONE.** When the child is seated at a table playing with a toy the experimenter names a pretext for leaving the room (which until the time of the experiment was unfamiliar to the child). The experimenter remains outside the room for two minutes. The child's behavior is recorded by concealed observation.

In this study the hypnotically regressed subject was left alone for five minutes, during which she was observed without her awareness.

2. **DARK ROOM.** While playing ball with the child the experimenter inadvertently throws the ball into a dark passage way, 18 feet long, leading from one corner of the room. The child is asked to retrieve the ball.

In the regression experiment the ball was thrown into an adjacent dark room about 21 feet long and 14 feet wide.

3. **STRANGE PERSON.** While the child was temporarily withdrawn from the room an assistant, dressed in a long gray coat, a large black hat, and a veil that obscured her features, seated herself in one of the two chairs near the entrance. When the child returned his reactions were observed when he noticed the stranger and when he was asked to obtain toys placed near the stranger's chair.

Apart from some variations in the clothing of the assistant, the hypnotic experiment was conducted in a similar setting.

4. **THE SNAKE.** A harmless garter snake, $22\frac{1}{2}$ inches long, was placed in a box deep enough so that it could not immediately climb out when the top was removed. In the box was placed a small colored toy. The child's attention was directed to the box, the lid was uncovered, the child was allowed to look in, if he voiced any questions the experimenter simply said, "It is a snake," and then pointed to the toy and asked the child to reach in and get the toy.

The hypnosis experiment utilized a garter snake 18 inches long.

Two additional fear situations were designed which are not reported on by Jersild. They involved a decapitated doll situation and a mouse experiment.

5. **DECAPITATED DOLL.** A "dydee doll" with a removable head is shown to the subject. She is permitted to hold it for a few minutes. The experimenter then takes the doll back and leaving the room for a moment removes the doll's head. He then re-enters the room and approaches the subject with the doll's head on his outstretched hand. The subject is confronted with the head directly and is urged to pick it up.

6. **MOUSE.** A small white mouse is placed in a cigar box. The subject is asked to open the box and bring the experimenter the contents.

The experimental situations were presented over a period of three weeks in the following sequence:

1. Being Left Alone
 - a.—Waking state
 - b.—Chronological hypnosis
 - c.—Hypnotic age regression to age 3
2. Dark Room, *a, b, c.*
3. Strange Person—*a, b, c.*
4. Snake—*a, b, c.*
5. Decapitated Doll—*a, b, c.*
6. Mouse—*a, b, c.*

The subject was a 25-year-old female well-trained in hypnosis and capable of post-hypnotic amnesia. Amnesia for each phase of the experiment was suggested and appeared to be operating.

C. RESULTS

Table 1 contains the results of this experiment and some objective findings from Jersild's studies (3). Of the four fear situations reproduced from Jersild's experiments, three produced fear in the hypnotic regression state. No fear was ever reflected in the waking or chronological hypnosis states. The "being left alone" situation failed to produce fear at any time. A reaction of fear was found by Jersild to occur in only 15.6 per cent of his study group in this particular situation and represents the smallest fear reaction of all the

TABLE 1
REACTIONS TO EXPERIMENTAL FEAR SITUATIONS BY CHILDREN AT EARLY AGE LEVELS
AND A SUBJECT IN HYPNOTIC AGE REGRESSION

Situation	N	Per cent showing fear	Jersild's data for children 36-47 months		Hypnotic age regression to age 3		Same subject in waking state of age 25		Same subject in hypnosis with no change in age orientation	
			Fear	No fear	Fear	No fear	Fear	No fear	Fear	No fear
Being left alone	45	15.6		X			X			X
Dark Room	45	51.1	X				X			X
Strange Person	45	22.2	X				X			X
Snake	36	55.6	X				X			X
Decapitated Doll	No Data		X				X			X
Mouse	No Data			X			X			X

situations included in his experiments. The "decapitated doll" produced fear in the regression state only. The "mouse" experiment failed to produce a reaction of fear in any state.

D. QUALITATIVE FINDINGS

1. **BEING LEFT ALONE.** The subject displayed no significant reactions to this situation in any of the experimental states. In hypnotic regression she seemed somewhat annoyed and was glad to see the experimenter upon his return to the room.

2. **DARK ROOM.** The only significant reaction took place in the hypnotic regression state. The subject upon entering the dark room cried out and immediately came back into the experimenter's room. She was trembling and on the verge of tears. Later, inquiry failed to elicit any explanation by the subject for her fear. She expressed much anxiety over the dark room and insisted upon keeping the door to it closed.

3. **STRANGE PERSON.** Again, only the hypnotic regression state produced any emotional reaction. Here the subject's reaction was initially closer to startle than fear. As she observed the strange person who made no overtures to her she became more frightened and began to cry.

4. **SNAKE.** The sight of the snake produced in the hypnotic regression state a paroxysm of terror. The subject screamed and cried. She was so upset as to require more than 20 minutes to calm down. In the waking and chronologic hypnosis states some disgust was expressed but no fear and the snake could be handled rather easily.

5. **DECAPITATED DOLL.** This stimulus produced by far the most intense expression of fear. Again, the fear was elicited in only the hypnotic regression state. The subject screamed and cried and did not immediately respond to direct suggestions for the termination of the reaction. She yelled, "Take it away, take it away." In addition to the extreme display of fear evidenced through crying and visual avoidance behavior it was later discovered that the subject had urinated in her pants during this experiment.

6. **MOUSE.** The mouse was handled with amusement in the regression state, with some annoyance when in chronological hypnosis and with some evidence of dislike in the waking state. No fear reaction was noted at any time in this situation.

E. DISCUSSION

It is of significance to note that whereas none of the situations which had been found to produce fear in young children produced fear in either the waking or hypnotic state for a 25-year-old adult, three such situations did

produce marked fear when the subject was placed in hypnotic age regression. One additional situation involving a decapitated head from a doll also produced extreme fear only in the hypnotic age regression state. The fear of a detached head is reported on by Hebb (2) in connection with spontaneous reactions in chimpanzees and elsewhere by the writer (5) in connection with hypnotically induced behavior. From the data revealed by this study it would appear that the effect of hypnotic age regression is in the direction of producing emotional reactions similar to those evidenced by children of an age level similar to that of the age regression. The spontaneity and intensity of the reaction of fear involved in these experimental instances including in one case involuntary urination appeared clinically to be beyond simulation. In addition to the intensity and nature of the subject's reactions there were post-experimental reactions which included the projection of fears in dreams and in some obsessive trends which will be reported on later but which now can be mentioned in adding to the impression of actuality of the emotional display.

Further work with developmental aspects of behavior in relation to hypnotic age regression will be necessary in order to fully evaluate the validity of this phenomenon and the evidence from this study would appear to warrant such further work. In addition to possible therapeutic value such a technique may offer a controlled approach to experimental studies of many of the factors involved in the psychogenetic and longitudinal aspects of personality growth. For if in experimental situations it would be possible to utilize the receptivity and reactivity set of childhood along with the cognitive, conceptual insight capacities of adult intelligence, intensive exploration of the origins of emotional behavior would be significantly enhanced.

F. SUMMARY

This paper has presented a study of reactions to established childhood fears situations by a subject in hypnotic age regression. The results generally revealed reactions of fear in the hypnotic age regression state only. No significant reactions of fear were noted in either the waking or chronological hypnosis states. From this brief experiment it would seem that in hypnotic age regression the subject's emotional reactions of fear assume a pattern similar to that of a child of that age. In this respect further evidence pointing toward the neuropsychological validity of hypnotic age regression is noted and the possibility of longitudinal psychogenetic studies of personality growth suggested. Further work in this area is currently in progress.

REFERENCES

1. BERGMAN, M. S., GRAHAM, H., & LEAVITT, H. C. Rorschach exploration of consecutive hypnotic chronological age level regressions. *Psychosomat. Med.*, 1947, **1**, 20-28.
2. HEBB, D. O. *The Organization of Behavior*. New York: Wiley, 1949. Pp. 335.
3. JERSILD, A. T. Studies of children's fears. *Child Devel. Monog.*, Teach. Coll., Columbia Univ., 1935. Pp. 360.
4. KLINE, M. V., & GUZE, H. The use of a projective drawing technique in the investigation of hypnotic age regression and progression. *Brit. J. Med. Hypnotism*, 1951, **3**, 1-12.
5. KLINE, M. V. A note on "primate-like" behavior induced through hypnosis: a case report. *J. Genet. Psychol.*, 1952, **81**, 125-131.
6. LEEDS, M. An hypnotic regression series. *Personna*, 1949, **1**, 13-16.
7. MERCER, M., & GIBSON, R. W. Rorschach content in hypnosis: chronological age level regression. *J. Clin. Psychol.*, 1950, **4**, 352-358.
8. SPIEGEL, H., SHOR, J., & FISHMAN, S. An hypnotic ablation technique for the study of personality development. *Psychosomat. Med.*, 1945, **5**, 273-278.
9. TRUE, R. M. Experimental control in hypnotic age regression studies. *Science*, 1949, 2866, 583-584.

Mental Hygiene Division
Department of Health
County of Westchester
White Plains, New York

THE WALL-SEEKING TENDENCY IN THREE INBRED MOUSE STRAINS (*MUS MUSCULUS*)*

Roscoe B. Jackson Memorial Laboratory

EMIL FREDERICSON¹

A. THE PROBLEM

The preference for walls often shown by rats and mice raised in the laboratory was found to lend itself to a comparative study of three mouse strains, each highly inbred.

The purpose of the present investigation consists specifically in a demonstration of hereditary differences in a behavior usually called the wall-seeking tendency. This behavior pattern is elicited by open spaces. Hall (1, 2) first suggested that the rat's preference for walls might be an interesting variable to study and used a standard piece of equipment to investigate it. Hall and Ballachey (4) have made extensive studies of this behavior in the rat. Patrick and Laughlin (6) have shown that this wall cathexis can be disrupted to a great extent in laboratory rats by means of large growing cages. Waters (9) investigated the influence of the wall-seeking tendency upon maze learning in white rats and found that more errors were made in the proximity of a wall. Lindzey (5) has demonstrated genetic differences in open-field emotionality, as measured by incidence of defecation, in five inbred mouse strains.

The present research, based upon the studies of Tryon (7, 8) and of Hall (3), is motivated by the need for further knowledge in regard to the influence of heredity upon behavior.

B. SUBJECTS

The subjects were 36 mice between five and seven months old, and represented the following three strains, each inbred by brother-sister matings for 30 generations; 12 C57 black mice of subline 10; 12 C3H mice, and 12 C Bagg Albinos. Males and females were represented in equal proportions. The mice lived as mated pairs and were kept in wooden cages 12 in. long, 6 in. wide, and 6 in. high.

*Received in the Editorial Office on March 4, 1951.

¹This investigation was supported by a research grant from the National Institute of Mental Health, U. S. Public Health Service. The writer wishes to thank Miss Jean R. Parker for invaluable assistance.

[illegible]

terms of the number of animals leaving the proximity of the wall and contacting or crossing the center square of the field. The strain differences are clearly revealed and indicate that the C57's are the least wall-seeking and the C Albinos the most. The C3H mice are closer to the C's than to the C57's. It should be noted that not a single C Albino mouse in the present sample, either male or female, ever left the proximity of the wall in order to cross through or even to contact the center part of the field.

Table 2 shows the mean number of center field contacts per mouse for

TABLE 2
AVERAGE NUMBER OF CENTER FIELD CONTACTS PER MOUSE
BASED ON TWENTY TRIALS

Strain	N	Average number of center field contacts for twenty trials
C57 black (10)	12	9.16
C3H agouti	12	.83
C Bagg Albino	12	.00

twenty trials. *T*-tests for differences in means of independent small samples are significant at well beyond the one per cent confidence level.

It is concluded that differences in the wall-seeking tendency of three inbred mouse strains exist and are most probably determined by hereditary variables.

F. DISCUSSION

The results of this study lend support to an emphasis upon the significance of heredity. Theories of learning and of personality have largely ignored the influence of heredity upon behavior. The unravelling of heredity-environment interactions is not an idle or hopeless task. Progress can be greatly accelerated by further studies of genetically controlled organisms. Relatively simple characteristics, such as wall-seeking, may be important determiners of more complex activities, e.g., learning, when these activities are measured in situations in which wall preference dominates the organism's behavior.

G. SUMMARY

Twelve C57 black mice, 12 C3H, and 12 C Bagg Albinos were tested in terms of their behavior in a field $2\frac{1}{2}$ ft. square, surrounded by a wall 1 ft. high. The results indicate a pronounced tendency in all three strains to stay close to the walls of the field. The incidence of leaving the wall and going to the center of the field was considerably higher among the C57 mice than either the C3H or Albino strains. The Albinos were never observed to leave

the proximity of the wall. Although the causes, hereditary or environmental, of this wall-seeking behavior are unknown, it may be permissible to infer the existence of genetically determined differences in a form of behavior which involves inhibitory processes.

REFERENCES

1. HALL, C. S. A comparative psychologist's approach to problems in abnormal psychology. *J. Abn. & Soc. Psychol.*, 1933, **28**, 1-5.
2. ———. Emotional behavior in the rat: III. The relationship between emotionality and ambulatory activity. *J. Comp. Psychol.*, 1936, **22**, 345-352.
3. ———. The genetics of behavior. In: *Handbook of Experimental Psychology*. (S. S. Stevens, ed.) (In press).
4. HALL, C. S., & BALLACHEY, E. L. A study of the rat's behavior in a field: A contribution to method in comparative psychology. *Univ. Calif. Publ. Psychol.*, 1932, **6**, 1-12.
5. LINDZEY, G. Emotionality and audiogenic seizure susceptibility in five inbred strains of mice. *J. Consult. Psychol.* (In press).
6. PATRICK, J. R., & LAUGHLIN, R. M. Is the wall-seeking tendency in the white rat an instinct? *J. Genet. Psychol.*, 1934, **44**, 378-389.
7. TRYON, R. C. The genetics of learning ability in rats—a preliminary report. *Univ. Calif. Publ. Psychol.*, 1929, **4**, 71-85.
8. ———. Genetic differences in maze-learning in rats. 39th Yearbook, National Society for the Study of Education, 1940, Part I, 111-119.
9. WATERS, R. H. The wall-seeking tendency and maze learning in the white rat. *J. of Psychol.*, 1937, **4**, 23-26.

Division of Behavior Studies

Roscoe B. Jackson Memorial Laboratory

Bar Harbor, Maine

BOOKS

The Journal of Genetic Psychology, the *Journal of General Psychology*, and the *Journal of Social Psychology*, will buy competent reviews at not less than \$2 per printed page and not more than \$3 per printed page, but not more than \$15.00 for a single review.

Conditions. Only those books that are listed below in this section are eligible for such reviews. In general, any book so listed contains one or more of the following traits: (a) Makes an important theoretical contribution; (b) consists largely of original experimental research; (c) has a creative or revolutionary influence in some special field or the entire field of psychology; (d) presents important techniques.

The books are listed approximately in order of receipt, and cover a period of not more than three years. A reviewer must possess the Ph.D. degree or its equal in training and experience.

Procedure. If among the books listed below there is one that seems important to you, you are invited to write a review of that book. It is not necessary to make arrangements with the Editor. Just send in your review. It does not matter if the book in question has been reviewed before.

1951

- McCLELLAND, D. C. *Personality*. New York: Sloane, 1951. Pp. 654.
- PASCAL, G. R., & SUTTELL, B. J. *The Bender-Gestalt Test*. New York: Grune & Stratton, 1951. Pp. 274.
- GUETZKOW, H., Ed. *Groups, Leadership, and Men*. Pittsburgh: Carnegie Press, 1951. Pp. 293.
- BRUN, R. *General Theory of Neuroses*. New York: International Univ. Press, 1951. Pp. 469.
- HELSON, H., Ed. *Theoretical Foundations of Psychology*. New York: Van Nostrand, 1951. Pp. 787.
- RAPAPORT, D. *Organization and Pathology of Thought*. New York: Columbia Univ. Press, 1951. Pp. 786.
- BERRIEN, F. K. *Comments and Cases on Human Relations*. New York: Harper, 1951. Pp. 500.
- VOLKART, E. H., Ed. *Social Behavior and Personality*. New York: Social Science Research Council, 1951. Pp. 338.
- CAMERON, N., & MAGARET, A. *Behavior Pathology*. Boston: Houghton-Mifflin, 1951. Pp. 338.
- GINZBERG, E., GINSBURG, S. W., AXELRAD, S., & HERMA, J. L. *Occupational Choice*. New York: Columbia Univ. Press, 1951. Pp. 271.
- KATONA, G. *Psychological Analysis of Economic Behavior*. New York: McGraw-Hill, 1951. Pp. 347.
- JAHODA, M., DEUTSCH, M., & COOK, S. W. *Research Methods in Social Relations*. New York: Dryden, 1951. Pp. 759.
- HARROWER, M. R., & STEINER, M. E. *Large Scale Rorschach Techniques*. (Sec. ed.) Springfield, Ill.: Thomas, 1951. Pp. 353.
- MILLER, G. A. *Language and Communication*. New York: McGraw-Hill, 1951. Pp. 298.
- SCHNEIDMAN, E. S., et al. *Thematic Test Analysis*. New York: Grune & Stratton, 1951. Pp. 320.
- HAYES, C. *The Ape in Our House*. New York: Harper, 1951. Pp. 247.
- WATSON, R. I. *The Clinical Method in Psychology*. New York: Harper, 1951. Pp. 779.

- JOSEPH, A., & MURRAY, V. F. *Chamorro and Carolinians of Saipan*. Cambridge: Harvard Univ. Press, 1951. Pp. 381.
- MARX, M. H., *Ed.* *Psychological Theory*. New York: Macmillan, 1951. Pp. 585.
- MEAD, M., & MACGREGOR, F. C. *Growth and Culture: A Photographic Study of Balinese Childhood*. New York: Putnam, 1951. Pp. 223.
- BUROS, O. K., *Ed.* *Statistical Methodology Reviews 1941-1950*. New York: Wiley, 1951. Pp. 457.
- TINBERGEN, N. *The Study of Instinct*. Oxford: Clarendon Press, 1951. Pp. 228.
- FREUD, A., *et al.* *The Psychoanalytic Study of the Child*, Vol. 6. New York: Internat. Univ. Press, 1951. Pp. 398.
- JEFFRESS, L. A., *Ed.* *Cerebral Mechanisms in Behavior*. New York: Wiley, 1951. Pp. 309.
- HUMPHREY, G. *Thinking: An Introduction to Its Experimental Psychology*. New York: Wiley, 1951. Pp. 331.
- 1952
- STONE, C. P., *Ed.* *Annual Review of Psychology*, Vol. 3. Stanford, Calif.: Annual Reviews, 1952. Pp. 462.
- HOCH, P. H. & ZUBIN, J., *Eds.* *Relation of Psychological Tests to Psychiatry*. New York: Grune & Stratton, 1952. Pp. 301.
- CATTELL, R. B. *Factor Analysis*. New York: Harper, 1952. Pp. 462.
- FOERSTER, H. VON, *Ed.* *Cybernetics*. New York: Macy Foundation, 1952. Pp. 240.
- BENDER, L. *Child Psychiatric Techniques*. Springfield, Ill.: 1952. Pp. 335.
- METTLER, F. A., *Ed.* *Psychosurgical Problems*. New York: Blakiston, 1952. Pp. 357.
- BENDA, C. E. *Developmental Disorders of Mentation and Cerebral Palsies*. New York: Grune & Stratton, 1952. Pp. 565.
- KINGET, G. M. *The Drawing-Completion Test*. New York: Grune & Stratton, 1952. Pp. 238.
- WOLFF, C. *The Hand in Psychological Diagnosis*. New York: Philosophical Library, 1952. Pp. 218.
- WOLFF, W. *The Dream: Mirror of Conscience*. New York: Grune & Stratton, 1952. Pp. 348.
- GORLOW, L., HOCH, E. L., & TELSCHOW, E. F. *The Nature of Nondirective Group Psychotherapy*. New York: Teach. Coll., Columbia Univ., 1952. Pp. 143.
- AUSUBEL, D. P. *Ego Development and the Personality Disorders*. New York: Grune & Stratton, 1952. Pp. 564.
- BYCHOWSKI, G., & DESPERT, J. L. *Specialized Techniques in Psychotherapy*. New York: Basic Books, 1952. Pp. 371.
- WHITE, R. W. *Lives in Progress: A Study of the Natural Growth of Personality*. New York: Dryden, 1952. Pp. 376.
- KARN, H. W., & GILMER, B. VON H. *Readings in Industrial and Business Psychology*. New York: McGraw-Hill, 1952. Pp. 476.
- ROBACK, A. A. *History of American Psychology*. New York: Library Pub., 1952. Pp. 426.
- KROEBER, A. L. *The Nature of Culture*. Chicago: Univ. Chicago Press, 1952. Pp. 437.
- FERGUSON, L. W. *Personality Measurement*. New York: McGraw-Hill, 1952. Pp. 457.
- BECK, S. J. *Rorschach's Test: III. Advances in Interpretation*. New York: Grune & Stratton, 1952. Pp. 301.
- VINACKE, W. E. *The Psychology of Thinking*. New York: McGraw-Hill, 1952. Pp. 392.
- KUHLEN, R. G., & THOMPSON, G. G., *Eds.* *Psychological Studies of Human Development*. New York: Appleton, Century, Crofts, 1952. Pp. 533.
- SNYDER, F. W., & PRONKO, N. H. *Vision with Spatial Inversion*. Wichita: Univ. Wichita Press, 1952. Pp. 144.
- JERSILD, A. T. *In Search of Self*. New York: Teach. Coll., Columbia Univ., 1952. Pp. 141.

CRITICAL REVIEWS OF RECENT BOOKS

The Journal of Genetic Psychology, 1953, 82, 149-150.

(Bühler, C., Smitter, F., Richardson, S., & Bradshaw, F. *Childhood Problems and the Teacher*. New York: Holt, 1952. Pp. 372.)

REVIEWED BY EDWIN S. SHNEIDMAN

This is an excellent book, a well written primer of dynamic psychology, for the practising elementary and high school teacher—it is not intended especially for the professional psychologist. The book is enlivened by an unusually large number (114) of well selected case illustrations, using subjects ranging in age from four to 17. The titles of the three parts of the book—there are 15 chapters subsumed under these parts—convey the breadth and content of the volume: (I) "A Psychological Approach to Children's Problems," (II) "What Can the Teacher Do?" and (III) "The Psychologist's Collaboration with the School," this last part including a well written (although elementary) chapter on projective techniques. The second part contains perhaps the heart of the book and deals with what the teacher can do in relation to scholastic problems, social problems, situational difficulties, remedial problems, parental problems, etc.

The book is written in easy to read, common-sense English. It moves along smoothly. Also, the format is good: there is thoughtful use of headings, illustrations, and various sizes and boldness of type to set off principles and case illustrations.

Occasionally, there is too brief discussion of a principle and occasionally a principle is illustrated with a case which could have served as an illustration for other (and even contrary) principles, but these do not constitute a major shortcoming in the text. There is much of Bühler's philosophy of development and maturation in this book; this consistent bias both adds considerably to the organization of the book and, at the same time, detracts somewhat from its scope and its usefulness.

In the Preface, the authors state their goals: "This book attempts to do three things: to give the teacher an understanding of the dynamics of behavior problems he may observe, to determine by means of case examples what the teacher may achieve in various circumstances, and to describe the type of problem with which the teacher needs specialized assistance." The authors

have succeeded quite well in achieving these goals and in producing an appealing volume in the bargain.

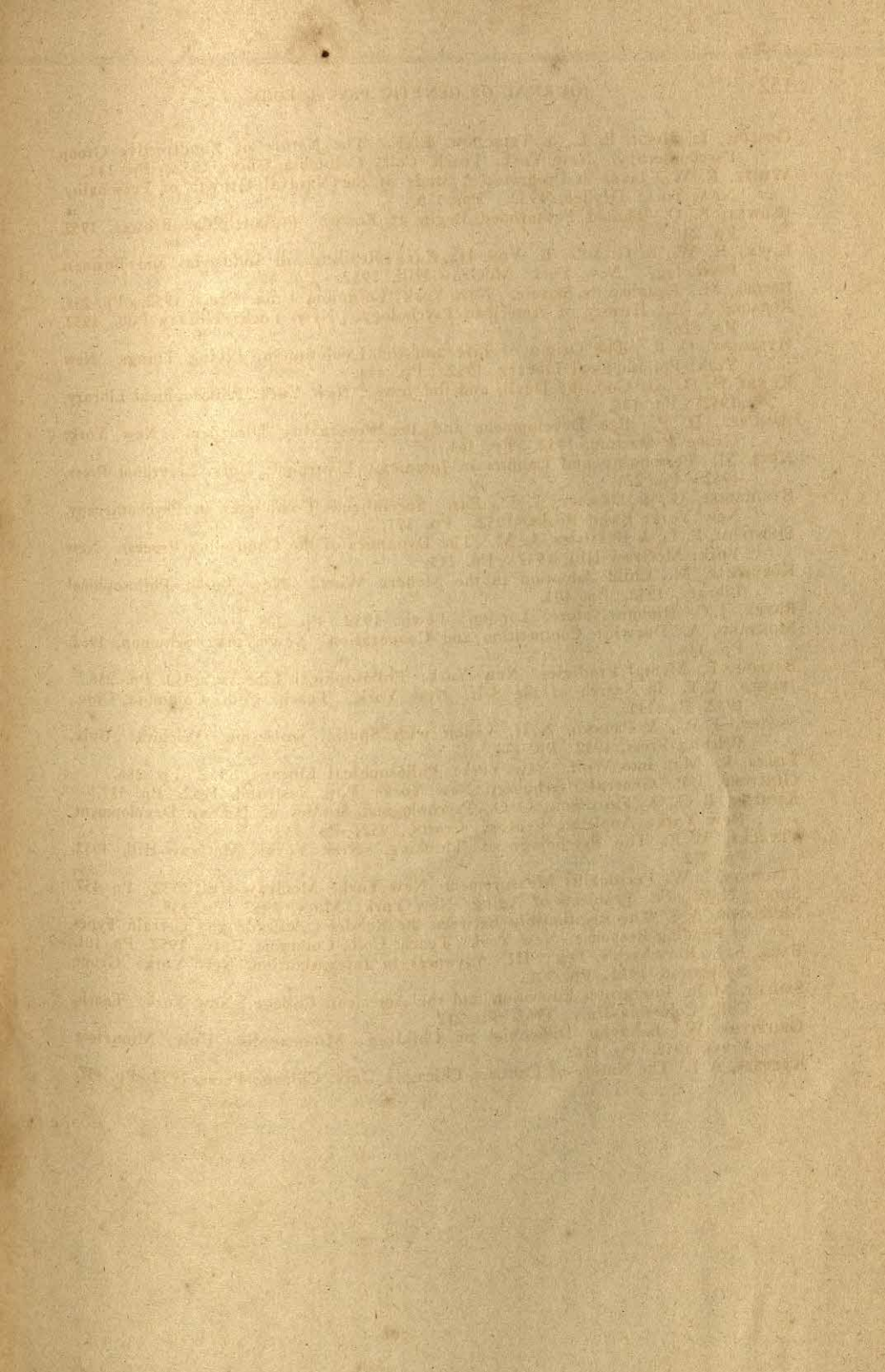
Veterans Administration Neuropsychiatric Hospital
Los Angeles 25, California

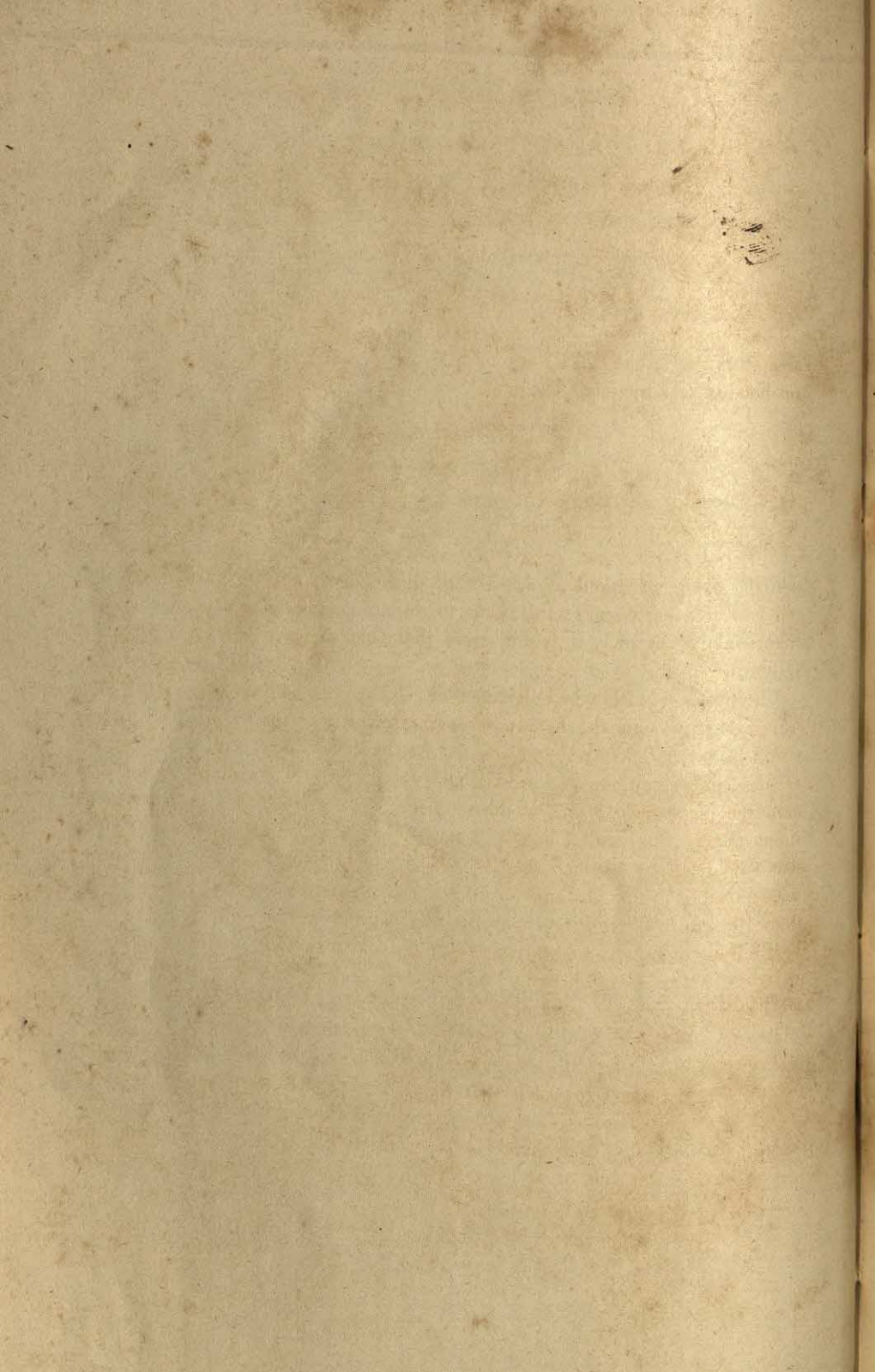
BOOKS RECENTLY RECEIVED

(There will always be two pages of book titles, listed in the order of receipt, i.e., the most recently received books will be found at the end of the list.)

- GREENACRE, P. Trauma, Growth, and Personality. New York: Norton, 1952. Pp. 328.
- PIGORS, P., & MYERS, C. A., Eds. Readings in Personnel Administration. New York: McGraw-Hill, 1952. Pp. 483.
- LAIRD, D. A. Increasing Personal Efficiency. New York: Harper, 1952. Pp. 294.
- BYCHOWSKI, G. Psychotherapy of Psychosis. New York: Grune & Stratton, 1952. Pp. 328.
- RUBIN, E. Experimenta Psychologica. Copenhagen: Munksgaard, 1949. Pp. 356.
- SMITH, M. B. The Single Woman of Today. New York: Philosophical Library, 1952. Pp. 130.
- CURRAN, C. A. Counseling in Catholic Life and Education. New York: Macmillan, 1952. Pp. 462.
- RAVEN, J. C. Controlled Projection for Children. London: Lewis, 1951. Pp. 176.
- BRYAN, A. I. The Public Librarian. New York: Columbia Univ. Press, 1952. Pp. 474.
- RUSSELL, B. Dictionary of Mind, Matter, and Morals. New York: Philosophical Library, 1952. Pp. 290.
- WOLFF, C. The Hand in Psychological Diagnosis. New York: Philosophical Library, 1952. Pp. 218.
- MENNINGER, K. A. A Manual for Psychiatric Case Study. New York: Grune & Stratton, 1952. Pp. 355.
- BENDA, C. E. Developmental Disorders of Mentation and Cerebral Palsies. New York: Grune & Stratton, 1952. Pp. 565.
- WOLFF, W. The Dream: Mirror of Conscience. New York: Grune & Stratton, 1952. Pp. 348.
- STAGNER, R., & KARWOSKI, T. F. Psychology. New York: McGraw-Hill, 1952. Pp. 582.
- FARIS, R. E. L. Social Psychology. New York: Ronald Press, 1952. Pp. 420.
- MIKESELL, W. H., & HANSON, G. Psychology of Adjustment. New York: Van Nostrand, 1952. Pp. 406.
- COUNTS, G. S. Education and American Civilization. New York: Teach. Coll., Columbia Univ., 1952. Pp. 491.
- BLOCH, H. A. Disorganization: Personal and Social. New York: Knopf, 1952. Pp. 608.
- HARTLEY, E. L., & HARTLEY, R. E. Fundamentals of Social Psychology. New York: Knopf, 1952. Pp. 740.
- CHESSER, E. Cruelty to Children. New York: Philosophical Library, 1952. Pp. 159.
- BENDER, L. Child Psychiatric Techniques. Springfield, Ill.: Thomas, 1952. Pp. 335.
- MUKERJEE, R. The Dynamics of Morals. London: Macmillan, (Undated). Pp. 530.
- METTLER, F. A., Ed. Psychological Problems. New York: Blakiston, 1952. Pp. 357.
- SCHIEDLINGER, S. Psychoanalysis and Group Behavior. New York: Norton, 1952. Pp. 245.
- GRAY, J. S. Psychology in Industry. New York: McGraw-Hill, 1952. Pp. 401.
- LEE, A. M. How to Understand Propaganda. New York: Rinehart, 1952. Pp. 281.
- CUSSLER, M., & DE GIVE, M. L. 'Twixt the Cup and the Lip. New York: Twayne, 1952. Pp. 262.
- CUSTANCE, J. Wisdom, Madness, and Folly. New York: Pellegrini & Cudahy, 1952. Pp. 254.
- BERRIEN, F. K. Practical Psychology. New York: Macmillan, 1952. Pp. 640.

- GORLOW, L., HOCH, E. L., & TELSCHOW, E. F. *The Nature of Nondirective Group Psychotherapy*. New York: Teach. Coll., Columbia Univ., 1952. Pp. 143.
- WHITE, R. W. *Lives in Progress: A Study of the Natural Growth of Personality*. New York: Dryden, 1952. Pp. 376.
- BROWER, E. O. *Should Fatherhood Begin at Forty?* Philadelphia: Brower, 1952. Pp. 240.
- KARN, H. W., & GILMER, B. VON H., *Eds.* *Readings in Industrial and Business Psychology*. New York: McGraw-Hill, 1952. Pp. 476.
- BERGER, M. *Equality by Statute*. New York: Columbia Univ. Press, 1952. Pp. 238.
- ROBACK, A. A. *History of American Psychology*. New York: Library Pub., 1952. Pp. 426.
- HYNDMAN, O. R. *The Origin of Life and the Evolution of Living Things*. New York: Philosophical Library, 1952. Pp. 648.
- RUNES, D. D. *Of God, the Devil, and the Jews*. New York: Philosophical Library, 1952. Pp. 186.
- AUSUBEL, D. P. *Ego Development and the Personality Disorders*. New York: Grune & Stratton, 1952. Pp. 564.
- KERR, M. *Personality and Conflict in Jamaica*. Liverpool: Univ. Liverpool Press, 1952. Pp. 221.
- BYCHOWSKI, G., & DESPERT, J. L., *Eds.* *Specialized Techniques in Psychotherapy*. New York: Basic Books, 1952. Pp. 371.
- SHOSTROM, E. L., & BRAMMER, L. M. *The Dynamics of the Counseling Process*. New York: McGraw-Hill, 1952. Pp. 213.
- KORNITZER, M. *Child Adoption in the Modern World*. New York: Philosophical Library, 1952. Pp. 403.
- RAVEN, J. C. *Human Nature*. London: Lewis, 1952. Pp. 226.
- MONTAGU, A. *Darwin: Competition and Cooperation*. New York: Schuman, 1952. Pp. 148.
- BARLOW, F. *Mental Prodigies*. New York: Philosophical Library, 1952. Pp. 256.
- JERSILD, A. T. *In Search of the Self*. New York: Teach. Coll., Columbia Univ., 1952. Pp. 141.
- SNYDER, F. W., & PRONKO, N. H. *Vision with Spatial Inversion*. Wichita: Univ. Wichita Press, 1952. Pp. 144.
- EISLER, R. *Man into Wolf*. New York: Philosophical Library, 1952. Pp. 286.
- GUILFORD, J. P. *General Psychology*. New York: Van Nostrand, 1952. Pp. 587.
- KUHLEN, R. G., & THOMPSON, G. G. *Psychological Studies of Human Development*. New York: Appleton, Century, Crofts, 1952. Pp. 533.
- VINACKE, W. E. *The Psychology of Thinking*. New York: McGraw-Hill, 1952. Pp. 392.
- FERGUSON, L. W. *Personality Measurement*. New York: McGraw-Hill, 1952. Pp. 457.
- SHOCK, N. W., *Ed.* *Problems of Aging*. New York: Macy, 1952. Pp. 138.
- McKILLIP, A. S. *The Relationship Between the Reader's Attitude and Certain Types of Reading Response*. New York: Teach. Coll., Columbia Univ., 1952. Pp. 101.
- BECK, S. J. *Rorschach's Test: III. Advances in Interpretation*. New York: Grune & Stratton, 1952. Pp. 301.
- SMILEY, M. B. *Intergroup Education and the American College*. New York: Teach. Coll., Columbia Univ., 1952. Pp. 212.
- GRIFFITHS, W. *Behavior Difficulties of Children*. Minneapolis: Univ. Minnesota Press, 1952. Pp. 116.
- KROEBER, A. L. *The Nature of Culture*. Chicago: Univ. Chicago Press, 1952. Pp. 437.





VISUAL DEPTH PERCEPTION IN THE RAT*

Department of Psychology, State College of Washington

ANN GREENHUT AND FRANCIS A. YOUNG

This investigation is concerned primarily with the ability of rats to perceive distance visually, and secondarily with the difference between albino and hooded rats in this respect.

A. PREVIOUS INVESTIGATIONS

The experimental designs used in the two experiments which are the direct precursors of the present employed a jumping situation, viz., the jumping of rats from one platform to another when the distance between them was altered at random. In a study by Richardson (6) the accuracy of the animal's jump was used as a mark of distance perception, since if the animal could make a successful jump to an arbitrarily selected and therefore unanticipated distance, it would seem that that distance had been perceived by him.

The problem with which Richardson was concerned was the sensory control mechanisms in the rat and her interest was primarily in the rôle of vision in the animal's adjustment, not in the animal's perception of distance as such. Her equipment consisted of two 12.5 by 12.5 cm. platforms each mounted on a 5 by 5 cm. wooden upright 50 cm. high, the uprights being supported by a horizontal base 120 cm. in length of 5 by 5 cm. timber, which was supported by wooden legs. One of the uprights was fixed, the other was movable. The distances between the platforms could be varied from 0 to 100 cm. The entire apparatus was painted white.

Training consisted of starting with the platforms very close together so that the animals could easily walk across and successively increasing the separation by one inch increments until they had to jump to cross to the other platform. A bit of food was placed on the landing platform as reward for each jump. During this series for each increase in distance the animals would scramble and for each shortened distance over one inch they would overjump, on succeeding jumps appearing to compensate for these and other errors. Five albino and three hooded animals finally learned to jump successfully to a distance of at least 55 cm. One albino, however, would not

*Received in the Editorial Office on May 2, 1951.

jump more than 37 cm., another learning the 55 cm. jump with difficulty.

Additional experimentation was performed using 55 cm. as a standard distance for which the animals were allowed to develop a habit. When a shorter distance was presented under this condition, the animals, as in the first series, overjumped, but if the decreased distance was much smaller than the standard one, 20 cm., for example, they could not adjust to it even on succeeding trials, although for greater distances they are reported to have adjusted "without great difficulty." Similarly when the distance was lengthened, the animals underjumped.

From these results, Richardson concluded that the rats used kinesthetic not visual cues in adjusting their jumps to various distances, accommodation being due to trial movements of the large muscles of the body. She did not report any difference between the two kinds of rats, albino and pigmented.

In a later investigation by Russell (7) a different and more refined criterion, capable of quantitative statement was used, namely the jump force exerted by an animal for various distances. Russell held that if an animal can perceive a given distance, he will alter the force of his jump accordingly. In other words, a positive correlation between amount of jump force and extent of distance to be jumped would be indication of distance perception, provided the distances had been presented in a random order to control for kinesthetic cues.

The experiment of Russell (7) was performed nearly 30 years later and was designed specifically around the problem of "depth perception in the rat." His apparatus, "... designed to measure the force with which a rat jumps from one ledge to another" (7, 137), consisted of two platforms, as in the Richardson study, one of which was supported on a pendulum that was sensitive to horizontal pressure on the platform. The displacement was read off a scale marked by a pointer at the lower end of the pendulum. Clamps attached to the pendulum permitted calibration. The jumping platform was 13 by 9 cm. in dimension with a narrow jumping ledge, 0.7 by 9 cm., elevated 1 cm. above the platform. The landing platform, which was moved along a centimeter scale, was 23 by 12 cm. with a landing ledge, 4.4 by 12 cm., elevated 4 cm. above it. At its far end bread and milk were placed. The equipment was of iron except for the platforms themselves.

His training procedure was similar to that of Richardson. His rats, deprived of food for 24 hours, were taught to cross from one platform to another to obtain food, the distance between the two platforms being gradually increased from 0 to 50 cm. The animals then jumped to distances presented

randomly before the test series. The latter consisted of six distances from 20 to 45 cm. with a 5 cm. interval, each distance being presented 10 times "in chance order." For each of the eight hooded and three albino rats used 10 jumps were run each day.

The results, presented as the total force used by each animal for each distance, show a direct relationship between force and distance for the hooded and, with some inversions, for the albinos. On the basis of the graph of these data Russell stated, ". . . it is clear that force and distance are directly related in the case of every rat. . . . In general, the curves are negatively accelerated" (7, 141). From the difference in the accuracy of their jumps compared with hooded rats and the lack of uniformity in the relationship between force and distance for the albinos, Russell concluded "that rats with non-pigmented eyes may be much inferior to rats with pigmented eyes" (7, 141).

For one hooded animal an additional series was run with the distances presented in successive instead of chance order. Russell noted that in this case there was a somewhat better discrimination. But the variability of these jumps was not presented. Russell concluded that the rat is able to perceive distance through vision alone and can adjust his neuro-muscular apparatus accordingly.

This experiment was performed to clarify the controversial results of the two experiments described above.

B. APPARATUS AND METHOD

The apparatus used was a modification of that of Russell's, designed to measure the vertical as well as the horizontal component of the jump force. Both jumping and landing apparatus were basically alike (see Figure 1). A horizontal bar, 32 cm. in length, was supported on ball bearings between uprights, 43 by 10 by 8.5 cm., which were mounted on a horizontal base, 30 by 16 cm. Attached to the bar were two arms, 5 cm. wide; one, 24 cm. in length, extended upwards at an angle of 75 degrees and was mounted with the jumping platform; the other, 43 cm. long, on the opposite side of the bar, extended downward at a 90 degree angle and was equipped at bottom with weights. On the platform side of the horizontal bar an arbitrary scale was attached by which the amount of displacement of the platform was measured. The platforms were placed on a table 64 cm. high, their bases in line with a meter stick fastened in the middle of the table, used for measuring the separation between them. Burlap was spread over the table to soften any falls which might occur. One platform was stationary, the other movable within a range from 0 to 70 cm. The entire equipment was made of unpainted wood.

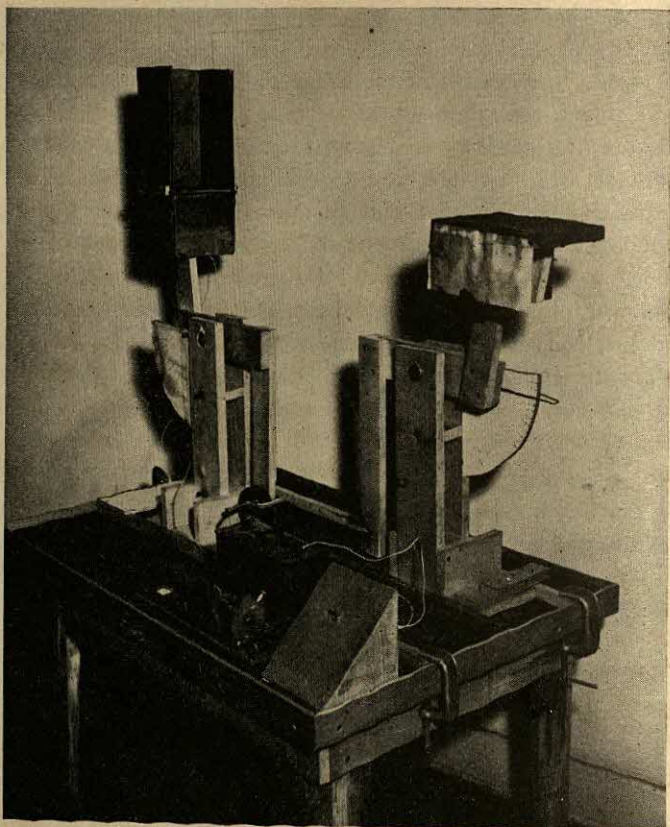


FIGURE 1
APPARATUS

Displacement of the jumping platform by the rat was actually due to the impulse of the jump, or change of momentum, rather than to the force alone. However, the arbitrary scale units (x) may be translated into the standard unit, grams of force (y), in terms of the amount of weight necessary to displace the platform a given amount when the weight is steadily applied. This relationship between weight and arbitrary units can be expressed in the following equation: $y = 10x + 10$. Thus, a 100 gm. weight, for example, would displace the platform nine arbitrary scale units. But weight, of course, is not equivalent to a change in momentum or to the "impulsive force." Results are, therefore, reported in terms of arbitrary

units. And when the term "jump force" is used, it should be remembered that reference is to the impulse of the jump.

It was necessary to add a metal shield to the jumping platform so that the animals would not climb off or jump to the wall. The lower part of the shield was 12.5 cm. long and extended completely around the platform. The upper part of the shield was 24 cm. high and covered three sides of the platform and the top. By this arrangement the animal could only get off the jumping platform by jumping in the general direction of the landing platform. The dimensions of the jumping platform were 14 by 13 cm. Since the lower arm held enough weight to balance the jumping platform, its appendages, and the heaviest animal used, under the platform there was a device for weight adjustments, in order to render the resistance uniformly proportionate to each animal's weight. Filled shotgun shells, ranging in weight from 10 to 90 gm., were used as the weight adjustments. On the floor of the jumping platform was an iron grid, wired to four dry cells of 1.5 volts each, a Harvard inductorium, and switch which were operated by the experimenter in an adjoining room.

A lower metal shield extending 12.5 cm. downward, was added to the landing platform, the dimensions of which were 23.5 cm. by 14 cm. Covering the floor of this platform was a thin layer of dark grey clay, so that the position of the animals' landing would be imprinted. An inch-scale for measurement of the accuracy of the jump was attached to one side of the platform. Three measures were thus taken for every jump, viz., the jump force, accuracy, and landing force. The last measure, however, could not be used since it was impossible to obtain a reading of the original landing force before the rebound of the platform altered it.

The walls of the room in which the experiment was run were covered with burlap to eliminate glare, but an empty sorting box containing a number of compartments was used to give a heterogeneous background to the rats' visual field. The experimenter worked in an adjacent room which was kept dim, but was not closed off from the room in which the rats made their jumps. The experimental room was dimly lit by a shaded 75-watt light globe, which was overhead and slightly to one side of the apparatus during the whole experiment.

The use of shock was an innovation to this type of experiment. Through a trial procedure during the training session it was found that a mild shock, irritating without being greatly disturbing, was most effective. A large shock would often upset the animal to such an extent that he would stop all activity, remaining "frozen" in a corner of the jumping apparatus. On

the other hand, a shock that was too mild seemed to have no effect on the animal. The Harvard inductorium was set at a reading of four to provide the most effective strength of the shock. Although great care was taken in the application of shock, the animals were emotionally disturbed throughout the experiment, as indicated by their excretory behavior and their resistance toward being put on the jumping platform.

Nineteen rats were used in the experiment, including eight hooded females, three hooded males, and eight albino males. All the animals were somewhat over 100 days old and naïve to the particular situation. Not included in the above total were four male albinos which had to be discarded at various stages of the training procedure because they simply would not jump. Possible explanations for this would be previous falls, which no animal entirely avoided, or indiscriminate shocking.

The animals were put through a training procedure like that of Richardson's and Russell's, the only difference being that they were learning an "avoidance response" from shock rather than an "approach response" to food. On an average an animal had from 20 to 40 trials in the training stage, starting with separations small enough to be walked across, which were increased by small intervals of 2.5 and 5 cm. out to the 50 cm. separation. Shock was used as the impetus for the jump and if the animal was preparing to jump he was not disturbed by being shocked. The animals were thus allowed to coordinate for their jumps and not driven from the jumping platform by a barrage of steady shocks.

The test series was composed of 70 jumps, 10 times to the following seven distances: 20, 25, 30, 35, 40, 45, and 50 cm. The entire 70 jumps were randomized together by means of Tippett's Table of Random Numbers. By such a method, one distance often appeared twice in a row, though no distance was ever to appear three times in succession. The same order of presentation was given to all the animals. Ten trials, or jumps, were given each night.

In addition to the test series with the random presentation of the distances, a control series was afterwards run with the distances presented in an ascending, successive order. Except for this change in the order of presentation, the situation was identical with the test series described above. Five albino males and five hooded females, selected from those who had participated in the first part of the experiment, jumped to each distance six times, starting with the 20 cm. separation. Again shock was used as the motivation and readings of the jump force and accuracy of the jump were taken.

C. RESULTS

The data presented in the following tables and graphs fall into three main parts: first, the practice series in which the animals were trained to jump by gradually increasing the separation of the platforms from 0 to 50 cm.; second, the test series, consisting of the 70 jumps with random order of presentation of the seven distances; and third, the control series in which the distances were presented in successive order from 20 through 50 cm. For each of these main parts both jump force and accuracy data are presented. In the case of the data on jump force, the 19 animals have been divided into three groups, eight albino males, eight hooded females, and three hooded males, except in the control series in which none of the latter animals were used. It was necessary to keep the two groups of hooded rats separate since they differed markedly in the amount of force exerted in jumping. However, for the data on accuracy the animals have been divided into only two groups, the 8 male albinos and 11 hooded, since no difference between the males and females of the latter appeared in this respect.

The average jump force exerted by each animal and the three group averages for each distance in the practice series are presented in Table 1. It should be emphasized that no effort was made to have the animals jump a given number of times during the training period, but rather merely to have each animal jump to each distance through 50 cm. in preparation for the test series. Thus, the number of jumps (N_j) to the various distances differs widely for any one animal and the total number of jumps differs among the different animals, as can be seen from the table. However, the data are still useful for comparison purposes. It can be noted that generally the greater the distance the more displacement was registered, without much difference appearing between the groups.

Table 2 gives the accompanying accuracy data from the practice series, presented for the two groups, 8 albino and 11 hooded, in terms of the proportion of successful and inaccurate jumps made to each distance. The inaccurate jumps have been broken down into scrambled jumps, in which the animal's hind legs did not land on the platform, missed or underjumps, and overjumps. Since the landing platform was 24 cm. long, the animals seldom went completely over the platform even though they had evidently overjumped. Thus, an arbitrary point demarcating accurate jumps from overjumps was determined on the basis of observation. Since the animals' lengths averaged approximately 16 cm., if their hind feet landed further than 5 cm. from the landing edge, they were considered to have overjumped. It is evident from the table that as the distances increase the proportion

TABLE 1
AVERAGE JUMP FORCE PER DISTANCE IN THE PRACTICE SERIES

Distance	1*	2	3	4	5	6	7	8	Total
<i>Albinos</i>									
20	2.50	3.25	1.30	1.75	3.00	2.66	3.30	1.75	2.44
25	2.40	3.18	3.50	3.50	3.30	3.25	4.00	1.75	3.11
30	3.90	5.40	5.25	6.25	6.00	4.00	4.75	3.75	4.79
35	5.00	7.00	4.50	7.30	8.00	6.25	6.66	6.25	6.40
40	6.96	7.07	5.00	7.25	7.12	7.33	7.50	6.75	6.87
45	9.50	7.00	6.25	10.00	9.50	9.87	8.25	8.00	8.55
50	8.62	7.87	8.75	9.08	10.00	9.75	9.10	7.66	8.85
N_j^{**}	50	38	34	27	34	39	31	29	
<i>Hooded Females</i>									
20	.50	.75	.37	1.25	.62	1.50	.50	.62	.76
25	1.37	2.00	.62	2.00	1.00	1.25	1.08	.83	1.27
30	3.41	3.75	2.00	2.00	1.91	3.15	2.41	2.70	2.67
35	3.66	3.50	3.25	3.33	2.08	4.50	3.17	3.81	3.40
40	6.56	5.45	5.08	7.41	5.00	6.50	6.00	5.91	5.99
45	7.50	5.55	5.50	8.50	5.00	7.25	6.60	5.41	6.41
50	7.25	6.65	6.37	9.25	5.62	8.25	7.87	5.91	7.15
N_j	16	25	18	19	20	23	23	20	
<i>Hooded Males</i>									
20	1.50	.75	2.00						1.42
25	3.50	2.08	2.87						2.82
30	5.12	3.05	4.00						4.29
35	5.62	4.25	6.25						5.37
40	8.08	7.37	8.62						8.02
45	9.70	7.68	10.00						9.13
50	11.00	9.00	9.37						9.79
N_j	16	20	19						

*The numbers refer to the individual animals.

**Total number of jumps made by each animal.

of accurate jumps decreases. There is one reversal for the hooded group at 40 cm., which has 10 per cent of the jumps more accurate than the preceding distance and no jumps which completely fell short of the landing platform. There seems to be a small difference between the two groups, the hooded animals having a greater proportion of accurate jumps. This difference is more marked as the distance increases.

The average force exerted in the 10 jumps to each distance of the test series is presented in Table 3 for the three groups. For the albinos there seems to be a rough relationship between force and distance for the first three distances only. Although there are several reversals for the individual hooded animals, there is a general trend for greater force to be used at greater distances. This is particularly noticeable in the two group averages for the hooded. The one reversal occurs at 40 cm., where the average force of

TABLE 2
PERCENTAGE OF ACCURATE JUMPS PER DISTANCE FOR THE PRACTICE SERIES

Distance	Successful jumps	Percentage of Scrambled jumps	Under-jumps	Over-jumps	Number of jumps
<i>Albinos</i>					
20	81	19	0	0	48
25	46	54	0	0	46
30	42	50	8	0	36
35	27	59	14	0	22
40	11	63	26	0	27
45	9	82	9	0	11
50	0	73	27	0	22
<i>Hooded</i>					
20	92	5	0	3	38
25	72	22	3	3	32
30	64	36	0	0	31
35	32	61	5	2	38
40	42	58	0	0	31
45	29	33	38	0	24
50	20	44	36	0	25

TABLE 3
AVERAGE JUMP FORCE PER DISTANCE UNDER TEST CONDITION WITH RANDOM ORDER OF PRESENTATION OF DISTANCES

Distance	1*	2	3	4	5	6	7	8	Total
<i>Albinos</i>									
20	7.72	7.52	6.52	8.72	8.00	6.45	7.02	8.12	7.64
25	8.12	7.90	8.70	8.85	8.55	7.15	8.10	8.28	8.21
30	8.82	8.85	8.62	9.00	9.18	7.35	7.82	9.42	8.63
35	8.90	7.66	8.48	8.65	8.48	7.32	7.70	7.90	8.13
40	8.80	7.98	8.45	8.05	9.10	6.65	7.90	8.12	8.13
45	8.18	7.68	8.12	8.12	7.35	6.40	7.15	8.58	7.70
50	8.80	8.15	9.02	8.35	9.00	6.62	8.45	8.58	8.37
<i>Hooded Females</i>									
20	4.90	3.96	2.88	5.72	4.12	4.83	5.02	4.80	4.53
25	5.45	4.15	3.75	6.48	3.90	4.75	5.10	5.62	4.90
30	5.22	4.26	3.50	6.82	4.14	5.65	4.98	5.52	5.08
35	7.12	3.42	4.18	6.82	3.95	6.32	5.45	5.42	5.34
40	4.38	3.39	3.98	6.40	3.95	5.48	5.38	5.15	4.76
45	7.45	4.30	4.22	7.18	4.00	6.15	5.52	5.88	5.59
50	7.22	5.18	4.20	6.98	4.48	5.92	5.59	5.70	5.70
<i>Hooded Males</i>									
20	8.30	6.02	7.98						7.43
25	9.42	6.85	8.65						8.31
30	9.20	6.90	8.98						8.36
35	9.50	6.86	9.30						8.53
40	8.90	6.68	9.38						8.32
45	9.20	7.92	9.52						8.88
50	9.82	7.98	9.40						9.07

*The numbers refer to the individual animals.

4.76 for the hooded females is lower than that for all the preceding distances but 20 cm., and similarly the average force of 8.32 for the hooded males is smaller than that for the two preceding distances. It may be noted that within all three groups there is only a small range from the smallest to the largest average.

The standard deviations for the averages of Table 3 are presented in Table 4. For all three groups the middle distances seem to have the most

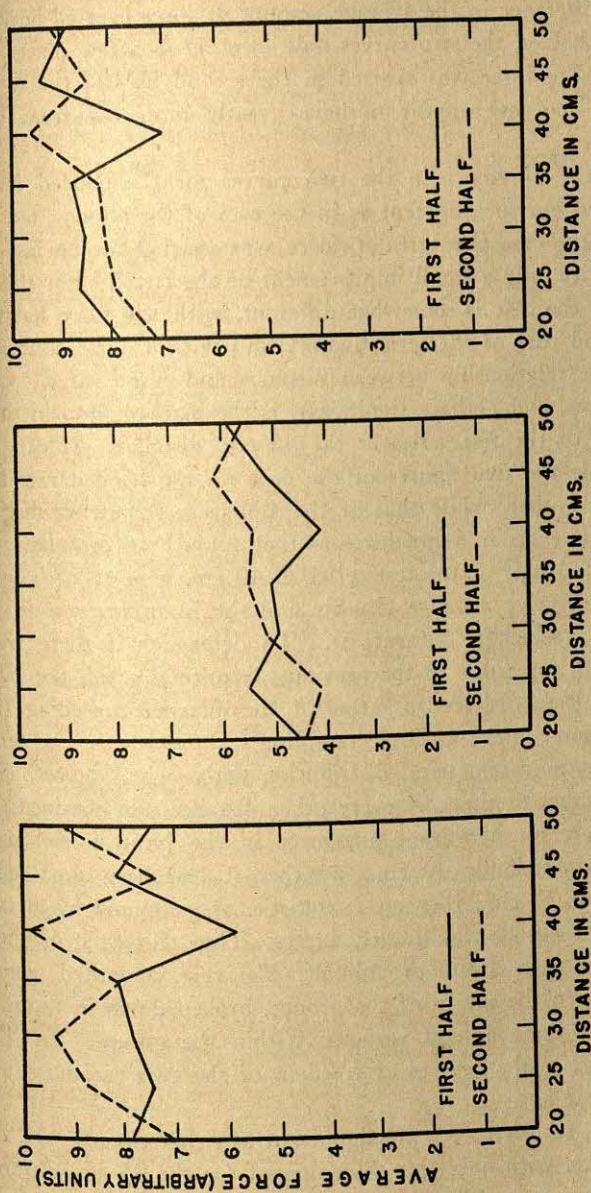
TABLE 4
STANDARD DEVIATIONS OF FORCE PER DISTANCE UNDER TEST CONDITION WITH RANDOM PRESENTATION OF DISTANCES

Distance	1*	2	3	4	5	6	7	8	Total
<i>Albinos</i>									
20	.80	.92	2.44	1.50	1.91	2.11	2.36	1.45	1.87
25	1.37	1.13	1.05	1.30	1.91	1.60	1.22	1.41	1.47
30	.81	1.30	1.59	1.85	2.17	1.39	2.26	1.46	1.81
35	1.57	1.22	2.00	1.64	2.14	1.27	1.36	1.46	1.73
40	1.52	1.76	2.40	2.01	2.78	2.01	2.27	2.21	2.27
45	2.46	1.94	1.64	2.85	2.53	2.49	2.42	1.52	2.38
50	2.38	1.17	.93	1.05	1.43	1.96	1.49	1.05	1.66
<i>Hooded Females</i>									
20	2.17	1.78	1.73	2.07	1.06	1.75	1.14	1.48	2.20
25	2.40	1.46	1.06	1.21	1.36	2.13	1.11	.95	1.78
30	2.92	2.36	1.66	1.97	1.79	1.31	1.38	1.18	2.09
35	2.06	2.34	1.13	1.50	1.60	1.04	1.48	1.59	2.09
40	1.51	2.07	1.41	1.62	1.74	1.56	1.61	1.08	1.86
45	1.73	2.01	1.06	1.46	1.56	1.32	1.23	1.36	1.96
50	1.57	.59	1.07	.93	.45	1.68	.55	.82	1.47
<i>Hooded Males</i>									
20	1.40	1.04	.71						1.48
25	1.62	1.19	1.46						1.78
30	1.16	1.72	1.26						1.74
35	1.37	1.10	1.07						1.72
40	1.50	1.54	1.70						1.97
45	1.32	1.55	1.34						1.56
50	1.63	1.10	1.32						1.56

*The numbers refer to the individual animals.

variability, although on the whole the sigma values remain fairly constant throughout the distances, instead of exhibiting any trend.

To determine if there were any important differences in performance on the test series due to practice, the jump force data were divided into two parts. One part consisted of the first five jumps, the second of the last five jumps to each of the seven distances. These data are presented for the three groups of animals in Figures 2, 3, and 4.



FIGURES 2, 3, & 4

RELATIONSHIP BETWEEN JUMP FORCE AND DISTANCE FOR THE FIRST AND SECOND HALVES OF THE TEST SERIES
 (Left, Figure 2, Albinos; center, Figure 3, Hooded Females; right, Figure 4, Hooded Males)

In the case of the albinos, shown in Figure 2, there can be little doubt that there is a difference in the average amount of force exerted between the first and second halves, the two curves only meeting in a few places. However, although more force was exerted in the second half of the test series, the relationship does not appear to have greatly improved from the first part.

In Figure 3 the difference in the two curves for the hooded females is neither so marked nor so consistent as in the case of the albinos, but in general it is again the case that greater force was exerted in the last half of the test series, with only a small improvement of the trend. For the hooded males, Figure 4, the case is somewhat different, with less force having been used in the second part of the series mainly, but with a more noticeable improvement in the relationship between distance and jump force. Common to all three figures, and perhaps significant, is the marked drop in the curve of the first half of the test series at the 40 cm. distance. It can be seen that on the whole the two figures of the data on the hooded rats bear resemblance to each other, while that of the albinos is somewhat distinct.

Such great differences in jump force as that found between the trends of the first and second half of the test series at 40 cm. would not be expected if the rat could perceive distance visually and the jump force were an adequate criterion by which to measure it. To determine if there were any effect on the jump force due to the previous distance jumped, the 10 jumps for each animal were ordered in terms of the distance preceding them in time of presentation. Since a simple sampling technique was used to obtain the random order, no symmetrical distribution with each distance being preceded a given number of times by every other distance was obtained. However, if there were no significant difference in the jump force used to a given distance, for example 40 cm., whether it had been preceded by a jump to 20 cm. or 50 cm., then an F statistic, as computed by an analysis of variance should not be significantly large. Thus, for each of the seven distances, and for each group of animals, F 's were computed, comparing the variance between the means of the groups, broken down in terms of the preceding distance jumped, and variance within the groups. In Table 5 are presented the F 's, the degrees of freedom of the two variances, and the significance of the F 's.

For the albinos, six of the F 's are significant at either the .05, .01, or .001 levels of confidence, with only that F computed for the 35 cm. distance not approaching significance. Thus, such large F 's would not be expected to occur more than the given percentage of times on the basis of sampling error,

TABLE 5

F STATISTIC SHOWING SIGNIFICANCE OF DIFFERENCES IN EFFECT OF PRECEDING DISTANCE ON THE JUMP FORCE

Distance	Degrees of freedom— Within groups	Degrees of freedom— Between groups	F	Significance level
<i>Albinos</i>				
20	75	4	3.20	Greater than .05
25	74	5	6.93	Greater than .001
30	75	4	16.34	Greater than .001
35	60	3	1.49	Less than .05
40	66	5	22.03	Greater than .001
45	59	4	15.75	Greater than .001
50	59	4	7.76	Greater than .001
<i>Hooded Females</i>				
20	75	4	2.10	Approaches .05
25	74	5	3.94	Greater than .01
30	75	4	5.39	Greater than .001
35	60	3	4.35	Greater than .01
40	66	5	7.74	Greater than .001
45	59	4	2.07	Approaches .05
50	59	4	1.76	Less than .05
<i>Hooded Males</i>				
20	25	4	3.25	Greater than .05
25	24	5	2.35	Approaches .05
30	25	4	1.61	Less than .05
35	20	3	—	—
40	22	5	3.31	Greater than .05
45	19	4	1.14	Less than .05
50	19	4	1.72	Less than .05

if the different groups were all from the same parent population. For the hooded females, four *F*'s meet the requirements of significance at one of the above levels of confidence, while two others, for the 20 and 45 cm. distances, would probably be significant between the .06 and .10 levels of confidence. Only the *F* for the 50 cm. distance does not approach significance. In the case of the hooded males, two *F*'s only, at the 20 and 40 cm. distances, are significant, while that for the 25 cm. distance approaches the .05 level of confidence.

Table 6, with regard to the accuracy of the jumps composing the test series, shows the proportion of accurate and inaccurate jumps for each distance. For the albinos, the accuracy of the jumps decreases sharply after the 30 cm. distance, where 82 per cent are accurate, to 51 per cent accurate at 35 cm. and only 14 per cent accurate, with 14 per cent misses, at the 40 cm. distance. The proportion of jumps which fell short of the landing plat-

TABLE 6

PERCENTAGE OF ACCURATE JUMPS PER DISTANCE UNDER TEST CONDITION WITH RANDOM ORDER OF PRESENTATION OF DISTANCES

Distance	Successful jumps	Percentage of		Over-jumps	Number of jumps
		Scrambled jumps	Under-jumps		
<i>Albinos</i>					
20	70	0	0	30	80
25	86	1	0	13	80
30	82	10	0	8	80
35	51	48	1	0	80
40	14	72	14	0	80
45	8	61	31	0	80
50	5	28	67	0	80
<i>Hooded</i>					
20	61	0	0	39	110
25	68	0	0	32	110
30	75	1	0	24	110
35	78	15	3	4	110
40	43	51	5	1	110
45	18	62	20	0	110
50	7	37	56	0	110

form reaches 67 per cent at 50 cm. There is a similar trend for the hooded, though the drop in accuracy comes after 35 cm. with 78 per cent accurate to 43 per cent at 40 cm. and 18 per cent at 45 cm. Fifty-six per cent of their jumps were inaccurate at 50 cm.

A comparison of the two groups indicates only slight differences, the general trend being the same. The hooded animals overjumped a somewhat greater per cent of the time at 20 cm. and much greater at 25 and 30 cm. than the albinos. However, the albinos had more underjumps at the greater distances than did the hooded.

To determine whether practice had any effect on the accuracy of the jumps, the test series was divided into five parts, each part consisting of two jumps to each of the seven distances. In Table 7 is presented the proportion of accurate and inaccurate jumps to each one-fifth part for the two groups. For the albinos there is a 16 per cent increase in accuracy from the first to the second part, which levels off for the remainder of the test series. In the case of the hooded group, the proportion of accurate jumps tends to increase through to the last part. A comparison of the two groups shows that there is no great difference in the proportion of accurate jumps until the last two parts, where the hooded improve in accuracy from 47 per cent in the third part to 58 per cent and 59 per cent in the fourth and fifth, while the albinos do not.

TABLE 7
PERCENTAGE OF ACCURATE JUMPS FOR EACH ONE-FIFTH PART OF THE TEST SERIES

Part*	Successful jumps	Percentage of Scrambled jumps	Under-jumps	Over-jumps	Number of jumps
<i>Albinos</i>					
I	32	31	23	14	112
II	48	33	14	5	112
III	47	25	21	7	112
IV	50	33	10	7	112
V	47	34	14	5	112
<i>Hooded</i>					
I	39	24	16	21	154
II	47	23	12	18	154
III	47	28	14	11	154
IV	58	24	6	12	154
V	59	19	11	11	154

*Each part consists of two jumps to each distance.

Of first importance is a consideration of whether the use of shock throughout the experiment has introduced a factor which in itself might have disturbed the performance of the animals. In order to determine what effect the shock had on jump force, the 10 jumps for each distance in the test series were broken down into those in which no shock had been used and into those in which it had for each animal. The t test of the significance of difference was then computed between these two distributions for the three groups at each distance. The t 's are presented in Table 8 along with the degrees of freedom (df) and significance.

Since some of the albinos had never jumped without shock to a distance, the t 's for the 35 and 45 cm. distances could not be computed for this group. Of the five t 's presented for the albinos, three are insignificant, at 20, 30, and 50 cm., while two are very significant, for 25 and 40 cm. For the hooded females only one of the seven t 's computed is significant, that for the 30 cm. distance, while none of the seven t 's on the data for the hooded males are significant.

Taking the three groups together, then, only three of the nineteen t 's were significant and two of these were based on the data of only three animals each. Of these three t 's no constant effect of the shock can be seen, the means with shock being greater than those without shock for the 25 cm. distance in the case of the albinos and for the 30 cm. distance in the case of the hooded females. With the other t , at the 40 cm. distance for the albinos, the mean on "no shock" jumps is greater than the mean for shocked jumps.

TABLE 8

t TEST SHOWING THE SIGNIFICANCE OF DIFFERENCES IN EFFECT OF SHOCK UPON JUMP FORCE UNDER TEST CONDITION OF RANDOM ORDER OF PRESENTATION OF DISTANCES

Distance	Degrees of freedom	<i>t</i>	Significance level
<i>Albinos</i>			
20	48	.50	
25	28	5.41	Greater than .001
30	38	.82	
35			
40	28	4.84	Greater than .001
45			
50	38	.02	
<i>Hooded Females</i>			
20	78	.15	
25	78	.39	
30	78	2.78	.006
35	78	.76	
40	78	1.70	.09
45	78	.78	
50	78	.56	
<i>Hooded Males</i>			
20	28	.49	
25	28	.28	
30	28	.07	
35	28	.48	
40	28	1.15	.12
45	28	.12	
50	28	.51	

To determine the effect of shock on accuracy it was necessary to hold distance constant, since more shock tended to be required at the greater distances and the latter tended to be related to inaccurate jumps (see Table 6). Thus, the test series was broken into two parts, the first composed of distances 20, 25, 30 cm.; the second, of distances 40, 45, 50 cm. These were the first two-fifths and last two-fifths distances, respectively. The amount of shock had been recorded only in general terms of 0, 5, 10, etc., shocks for each jump, so only three categories of number of shocks were used: 0 shock, 1-10 shocks, and 11 or more shocks.

Table 9 gives the proportion of accurate to inaccurate jumps for the three amounts of shock for both the first three and last three distances. N_j stands for the total number of jumps under each condition. It is clear from the table that for both the albinos and hooded the proportion of accurate jumps is greater under all three amounts of shock for the smaller distances, 20, 25, and 30 cm., and smaller for the larger distances, 40, 45, and 50 cm. This is

TABLE 9
PERCENTAGE OF ACCURATE TO INACCURATE JUMPS FOR THREE AMOUNTS OF SHOCK,
HOLDING DISTANCE CONSTANT

Amount of shock	Distances 20-30			Distances 40-50		
	Accurate	Inaccurate	N_i^*	Accurate	Inaccurate	N_i
<i>Albinos</i>						
0	76	24	25	4	96	24
1-10	84	16	168	9	91	133
11 or more	89	22	47	6	94	83
<i>Hooded</i>						
0	70	30	212	20	80	141
1-10	66	34	116	24	76	171
11 or more	50	50	2	17	83	18

*Number of jumps.

consistent with the accuracy data presented in Table 6, showing the relationship between distance and accuracy. Interesting to note, as far as comparison of the two groups is concerned, is that the albinos jumped mainly under shock condition whereas the hooded jumped the greater proportion of times without shock.

To determine the relationship between force and distance when the rat is allowed to correct for his error and is not dependent solely upon visual cues, an additional series of jumps were given to 10 of the trained animals, five albinos and five hooded females. They were presented six times with each distance in successive order starting with 20 cm. Table 10 gives the average

TABLE 10
AVERAGE JUMP FORCE PER DISTANCE UNDER CONTROL CONDITION WITH SUCCESSIVE
ORDER OF PRESENTATION OF DISTANCES

Distance	1*	2	3	4	5	Total
<i>Albinos</i>						
20	7.54	9.12	6.12	7.54	9.66	7.90
25	8.25	8.00	7.54	7.33	8.46	7.91
30	9.71	9.46	8.83	8.41	10.08	9.30
35	9.25	8.59	8.54	9.21	10.12	9.14
40	10.66	9.88	10.25	10.17	12.33	10.67
45	12.50	11.00	11.25	12.37	15.41	12.51
50	14.08	12.16	12.00	13.46	17.66	13.87
<i>Hooded Females</i>						
20	4.83	4.08	3.83	4.83	4.75	4.47
25	5.58	4.33	4.08	5.58	5.29	4.98
30	6.96	5.54	4.71	6.41	6.08	5.94
35	9.16	6.87	7.12	8.75	7.25	7.83
40	10.25	7.12	7.41	9.08	7.83	8.34
45	11.29	8.54	8.33	9.66	8.75	9.31
50	12.16	9.33	9.79	10.50	9.16	10.19

*The numbers refer to the individual animals.

amount of force used by each animal for each distance and averages for the two groups in this control series.

In the case of the albinos, with only a few inversions, there is a general increase in force for an increase in distance, which becomes more pronounced for the larger distances. The inversions occur mainly at the 20 cm. and 35 cm. distances for the individual albinos. A consistent, uninverted trend is found in the case of every hooded animal. The greatest increase for this group is at 35 cm. Comparing the two groups, it is evident that the similarities outweigh the differences. The most striking discrepancy is at the 35 cm. distance, where the albinos have their one inversion in the trend while the hooded have their greatest increase.

Standard deviations for the average jump force of Table 10 are presented in Table 11. No trend can be seen in these data for greater varia-

TABLE 11
STANDARD DEVIATIONS OF FORCE PER DISTANCE UNDER CONTROL CONDITION OF
SUCCESSIVE PRESENTATION OF DISTANCES

Distance	1*	2	3	4	5	Total
<i>Albinos</i>						
20	.23	1.26	.59	.73	1.38	1.66
25	.48	.12	.86	.72	.48	.68
30	.93	1.02	.66	.85	.92	1.05
35	.96	.56	.57	.73	1.07	.99
40	.95	.85	.73	.84	.94	1.21
45	.82	.28	.82	.86	1.68	1.82
50	1.28	.75	.85	.22	.88	2.19
<i>Hooded Females</i>						
20	.87	.56	.64	.30	.24	.67
25	.49	.39	.26	.49	.24	1.18
30	.45	.39	.40	.30	.62	.90
35	.40	.10	.47	.48	.81	1.09
40	.20	.56	.20	.30	.52	1.26
45	.58	.50	.41	.30	.39	1.14
50	.40	.45	.73	.41	.10	1.20

*The numbers refer to the individual animals.

bility to accompany the greater distances. There are no striking group differences. Generally it can be noted that there is only small variation around the average force when the distances are presented successively.

The accuracy of the jumps in the control portion of the experiment is shown in Table 12 for the two groups. The albinos show a big drop in accuracy from the 20 cm. to the 25 cm. distance of from 90 per cent accurate to 73 per cent. This is followed by another marked decrease in accuracy at

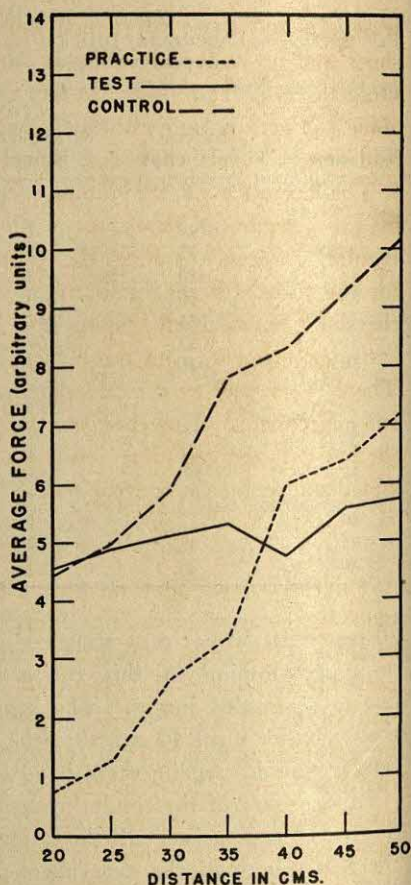
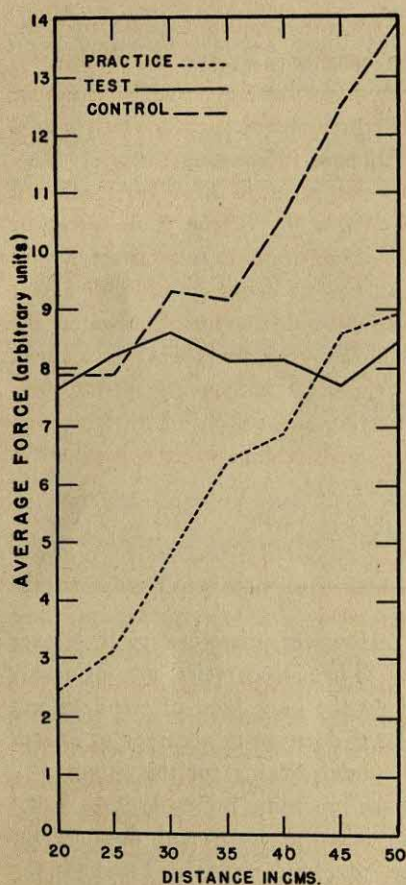
TABLE 12

PERCENTAGE OF ACCURATE JUMPS PER DISTANCE UNDER CONTROL CONDITION WITH SUCCESSIVE ORDER OF PRESENTATION OF DISTANCES

Distance	Success- ful jumps	Percentage of			Number of jumps
		Scrambled jumps	Under- jumps	Over- jumps	
<i>Albinos</i>					
20	90	7	0	3	30
25	73	27	0	0	30
30	50	50	0	0	30
35	60	40	0	0	30
40	53	47	0	0	30
45	33	64	3	0	30
50	10	83	7	0	30
<i>Hooded Females</i>					
20	100	0			30
25	97	3			30
30	80	20			30
35	83	17			30
40	57	43			30
45	47	53			30
50	20	80			30

30 cm. with 50 per cent accurate jumps. However, there is very little over or underjumping in this entire series. The inaccuracies are primarily due to scrambled jumps. The hooded show the same kind of trend, except that it is not until 40 cm. that the first large decrease in accuracy of 26 per cent is found. Again, there is no over or underjumping for this group.

In summary of the results presented in the preceding tables, Figures 5 and 6 give a comparison of the relationship between force and distance for the three parts of the experiment—practice, test, and control. For both the albino and hooded female groups, the three curves are distinct. Striking is the similarity in the curves of the practice and the control series, both of which had successive order of presentation of the distances. In the case of the test series, the curve falls in an intermediate position on the y-axis, but is of an entirely different shape. In comparison with the other two curves, it seems almost horizontal. The total change in this curve between the 20 and 50 cm. distances is not so great as that found between two successive distances of the other curves, 35 and 40 cm., for example. Consistent with Figures 2 and 3, the amount of force appears to increase with the amount of practice. This relationship is particularly clear for the middle distances, 25 to 45 cm. There is some overlapping of the test series curve with the other two for the extreme distances.



FIGURES 5 & 6

COMPARISON OF PRACTICE, TEST, AND CONTROL SERIES ON RELATIONSHIP BETWEEN JUMP FORCE AND DISTANCE

(Left, Figure 5, Albinos; right, Figure 6, Hooded Females)

D. DISCUSSION OF RESULTS

1. Effect of Shock

Before anything can be said about either the criteria of jump force and accuracy or of visual distance perception, it is necessary that there be no methodological variable which might be distorting the results. Such a variable might have been the shock motivation in the present experiment. The t 's computed for each distance between these jumps for which shock

was used and those for which shock was not used would indicate that the shock had no noticeable effect on the jump force used by the rat (Table 8). Only three of the t 's were significant which might have resulted from chance alone. There is no consistency among the three significant t 's which would lead one to expect that some common factor was influencing them.

The accuracy of the jumps, also, does not seem to have been affected by the amount of shock, as is shown in Table 9. There is no great or regular difference among the proportion of accurate to inaccurate jumps for the three different amounts of shock. It is, of course, a possibility that the shock introduced a disturbing and variable effect which permeated the entire experiment, even those parts in which the shock itself was not applied. There is no way of directly determining whether this happened in the present experiment. However, evidence is offered against such an hypothesis on the basis of the complete consistency of the relationship between jump force and distance in the control series (see Table 11).

2. *Jump Force Used in Test Series*

Table 4 shows that no clear relationship exists between force and distance in the test series. For individual rats there are many inversions, some of these disappearing when the data are grouped. For the hooded animals particularly, the group data yield a slight trend except for the one inversion at the 40 cm. distance. In the case of the albinos there is a tendency for greater force to be used at greater distance up to 35 cm., after which fluctuations occur. Accepting this at face value, it could possibly be concluded, on the basis of the jump force criterion, that visual distance perception exists in rats and that it is more extensive in the hooded than in the albino. This is exactly the conclusion that Russell (7) arrived at on the basis of his data, which were far more regular. However, this would seem a hasty generalization on the basis of the present data. No individual rat completely conformed to the group data and the latter only suggest a trend. Perhaps the relationship found is merely an artifact of this particular order of presentation of the distances and the combining of data.

To check this hypothesis, the F 's presented in Table 6 were computed for each distance, the jumps categorized in terms of the preceding distance jumped. In the case of the albinos the F 's indicate that there is a significant difference in the jump force used to a given distance depending upon the distance jumped preceding it. Similarly for the hooded females, it can be seen that the preceding distance has an effect upon the jump force to any given distance, four of the F 's being significant at least the .05 level of

confidence and two of them at greater than the .10 level. The data presented on the hooded males do not give such clear-cut evidence. Only three of the seven F 's can be considered significant, those for the 20, 25, and 40 cm. distances. Yet it would seem unparsimonious to infer that in the special case of the three hooded males, who otherwise do not differ from the hooded females, there was no effect of the preceding distance on the jump force at the four distances, 30, 35, 45, and 50 cm., while at the other three distances there was. Preferably, it could be suggested that perhaps the differences did not appear for these distances due to the small number of hooded males, and that pending further investigation, the data for the hooded males be interpreted as the data of the two larger groups. Supporting such an interpretation is the general tendency for the group means to increase as the preceding distance in all three groups.

In summary, then, it would appear that with the order of presentation of distances randomized, the animals could not make the proper adjustments so that the jump force expended would be appropriate to the change of distance. But before this can be interpreted as an absence of visual distance perception in the rat, the jump force criterion must be evaluated. This will be done after the discussion comparing the test and control series.

3. Accuracy of Jumps in the Test Series

In respect to the accuracy of the jumps under test condition, Table 6 shows a great drop in accuracy for the larger distance with an increasing proportion of underjumps, the animals being most accurate to the smaller distances. These data might be interpreted *prima facie* as indicating that the animals were not able to perceive the larger distances since over half of their jumps to them were inaccurate. It is at least obvious that there is a difference in the proportion of accurate jumps between 25-35 cm. as compared with 40-50 cm. This difference must await interpretation until after the accuracy of the jumps in the control series is discussed.

After some practice the hooded rats made a greater proportion of accurate jumps, as is shown in Table 7. The albinos apparently benefited by practice only up to the second part of the experiment. This might be an indication that there should be a longer period of practice with an introduction of random order of presentation of distances before the test series proper is commenced.

4. Comparison of Practice, Test, and Control Series.

Figures 5 and 6 indicate that a significant difference obtains in the relationship between jump force and distance depending upon the order of

presentation of the distances. The curves of the practice and control series show a steady rise, while the curve of the test series remains almost horizontal, even for the hooded animals. Practice appears to have increased the amount of force exerted but in itself has had little or no effect on the relationship here discussed. From the former curves it can be seen that jump force, as measured by the equipment used in this experiment, is at least a potentially adequate criterion since a strong relationship between force and distance has been obtained and that the range of distances used did not approach the physiological limits of the rats' jumping capacity since the curves of the control series are not negatively accelerated at 50 cm.

With respect to the variability of the jumps, Tables 5 and 12 show again a difference for test and control. The standard deviations are in all cases smaller for the hooded females under the control condition and mainly so for the albino. This supports, perhaps, the above conclusion based on the F 's that the jump force at any given time is partly determined by the distance previously jumped, since in the control series the same distance followed itself for five times in a row whereas for the test series at no time did the same distance follow itself more than once.

The accuracy of jumps also differs under test and control conditions. The advantage of a successive order of presentation is mainly at the two extremes, 20 and 50 cm., where for the most part under- and overjumps do not occur in the control series. The practice series shows the smallest proportion of accurate jumps which would be expected since the animals were just learning how to jump and becoming acquainted with the various distances.

No striking difference is found between accuracy for the test and control series. In both cases a sharp decline in accuracy occurs with an increase of distance. This would suggest that accuracy alone is not to be taken as absolute indication of visual distance perception. Since even in the control series, where other cues were available, only a small proportion of accurate jumps were made to the larger distances.

5. *Evaluation of the Criteria*

On the basis of the foregoing discussion it is now possible to give an evaluation of the criteria used. If the two criteria are found to be adequate, then it must be concluded that rats do not possess visual distance perception, because a strong relationship between jump force and distance appears only when the latter is presented in a successive order, permitting kinesthetic cues and correction for poor jumps. Furthermore, it is only in the latter

case that the animals avoid short and overjumps. Thus, it is important to evaluate the criteria.

From the data of the present experiment specific instances of inconsistencies of the jump force criterion can be given. One such inconsistency was that the force used for a set of jumps did not conform to the common-sense assumption that greater force propels the animal further. For example, the following data were obtained from hooded Male 2 for Jumps 55 to 59 (Table 13), where the figures under the accuracy column refer to the number of inches from the edge of the landing platform that the hind feet of the animal landed. Here the animal landed at a greater distance from the edge of the platform at 40 cm. than at 25 cm., even though less force was used in the former case.

TABLE 13

Distance	Force	Accuracy
40	7.50	$\frac{3}{4}$
25	8	0
40	7.25	$\frac{1}{2}$
30	7.75	$\frac{1}{2}$
40	7.50	$\frac{1}{2}$

Again, the same force was used for a distance at two different times, the jump being accurate at one time but not the other. This is exemplified in the case of hooded Female 2 for Jumps 55 and 57, wherein both were to 40 cm., with the same amount of force of four units used. Yet one of the jumps was scrambled, the other accurate.

A third type of inconsistency may be illustrated by the following data on hooded Female 1, where the same force was used to two different distances, with accuracy at the greater distance and scrambling at the smaller (Table 14).

TABLE 14

Distance	Force	Accuracy
50	7	0
40	4.5	1
45	7	$-\frac{1}{2}$

These examples are sufficient to show that the force used to propel the animal a given distance varies, depending upon factors independent of the perception of the distance of the two platforms. Thus, jump force cannot be considered a totally satisfactory criterion, since it has been shown that an animal can gain a given distance with varying amounts of force.

However, there is reason to believe that the jump force criterion need not be discarded entirely but only that it be used and interpreted with caution. It can be seen from the control series (Figures 5 and 6) that a correlation does exist between jump force and distance, yet the only change that was made from the test situation was the successive order of presentation of distances. Thus, since when the animal is allowed additional cues to the visual he uses a consistent amount of force, but when visual cues alone are available he is unable to make the appropriate modifications, it would still seem justifiable to assume that the rat does not possess visual distance perception at all or only rudimentarily.

As has been previously observed, accuracy does not suffice as a single criterion of visual distance perception either, since the animals did not jump with marked difference in proportion of successful jumps when the distances were presented in successive order, where corrections were possible, than when the distances were presented in random order. However, there is a difference in the cases of overjumps and underjumps, which would suggest that accuracy might be used as a supplement to the jump force criterion. Certainly the accuracy of the jump cannot be ignored, as it has been by Russell (7), since it is only on the assumption that the animal is attempting to reach the other platform that the jump force criterion can be used. Perhaps some way of combining these two criteria could be developed so that the deficiencies of the one are met by the advantages of the other.

6. *Comparison of Hooded and Albino Animals*

The differences between albino and hooded rats do not appear as pointed as that which Russell (7) reported. Far more striking is their similarity of performance in both the accuracy and the amount of force of the jumps for the three parts of the experiment. Such differences as do exist, however, do tend to favor the hooded. Thus, the latter show a greater proportion of accurate jumps and a smaller proportion of underjumps in the test series (see Table 7). The hooded rats also seem to have improved in accuracy through the greater part of the test series while the albinos made no improvement after the first few jumps (Table 8). Moreover, the hooded animals tended to show a slightly better relationship between jump force and distance in the test series (Table 3). These differences are slight, however, and are possibly due to other physiological conditions which differ in the two breeds, rather than to the absence of pigmentation in the albinos alone, as Russell suggests (7). There is no way of determining this on the basis of the present experiment, but it is advisable that further investigation

be conducted before such statements as Russell's be accepted. Particularly is this advisable since even in the control series where the visual cues were supplemented by kinesthetic cues and learning, the pigmented animals, i.e., the hooded females, still maintained their superiority over the albinos in proportion of accurate jumps, and in consistency of trend between force and distance (Figure 6).

7. *Comparison between Present and Russell's Data*

It is difficult to explain why there should be such great discrepancies between this experiment and Russell's (7). Yet Russell reports the following results with which this experiment disagrees:

1. The existence of a strong relationship between jump force and distance when the distances were presented in random order.
2. A tendency for the variability of the jump force to increase as the size of the distance.
3. A relatively small amount of improvement in the relationship between jump force and distance when the distances were presented in successive rather than in random order.
4. A striking difference in performance between pigmented and albino rats in the relationship of force to distance and in the accuracy of the jumps, in both cases the hooded rats being far superior.

The two main methodological differences between the experiments are in motivation and equipment. But examination of shock for variable effects has shown that it has had no noticeable effect on the animals' performances (Tables 8 and 9). Russell does not present complete information regarding the calibration of his equipment but there is no reason to believe that it would cause such differences as the above. If it is correct that the preceding distance affects the jump force to a given distance, then the order of presentation of the distances is vital information, but Russell does not elaborate on this point. If Russell randomized the order of presentation of each distance but ran 10 consecutive trials to each distance before presenting the next one [Shaad's and Helson's psychophysical variations (8)], then his test series is comparable to the control series of the present experiment, both of which show an increase in jumping force with increase in distance. The difference between his test curve, which is negatively accelerated, and the control curve obtained in the present experiment, which is almost linear, can be explained in terms of differences in equipment. That is, Russell's apparatus did not measure the vertical component of the jump force which becomes more and more important as the distance to be jumped is increased.

E. CONCLUSIONS

In summary the following points may be noted:

1. The relationship between jump force and distance is negligible when the order of presentation of the distances is randomized (Table 3).
2. The amount of force used for a particular jump is in part dependent upon the preceding distance jumped (Table 5).
3. The accuracy of the jumps in which order of presentation of the distances is random is at an optimum at the intermediate distances, overjumps occurring at the small distances and short jumps occurring at the larger ones (Table 6).
4. Practice is directly related to the amount of force exerted, but only slightly improves the force-distance relationship (Figures 2, 3, 5, 6).
5. Practice improves the accuracy of the jumps, within certain limits, which were reached early in the experiment by the albinos, but only towards the last by the hooded (Table 7). On the basis of points four and five, it is suggested that the practice period should have been extended in length and should have also included distances presented in random, as well as successive, order.
6. Shock has no particular effect on either the amount of jump force or the accuracy of the jump (Tables 9, 10).
7. There is a great difference in the relationship between jump force and distance, depending upon the order of presentation of the distances; the relationship is evident only when a successive order of presentation of distances is used.
8. Accuracy to the extreme distances improves in terms of smaller proportion of short and overjumps when the distances are presented in successive order as compared with random order of presentation, but does not improve greatly in terms of proportion of successful jumps to the greater distances.
9. The similarities in performance of the pigmented and non-pigmented rats in terms of both accuracy and jump force are far more outstanding than the small differences, but such differences as do exist tend to favor the pigmented animals.
10. The adequacy of the jump force criterion is open to the following criticism. If the jump force criterion were adequate, one would expect that for any animal there would be a constant relationship between the amount of force and the distance gained by the animal, but the following inconsistencies were found: (a) The same force was used but different distances were gained by the animal, (b) different amounts of force were used but the same dis-

tance was gained by the animal, and (c) different amounts of force were used for different distances gained but the greater amounts of force were used with lesser distances gained and the lesser amounts of force were used with greater distances gained.

11. The adequacy of the accuracy criterion is in certain respects also open to criticism. Assuming the accuracy criterion were adequate, unsuccessful jumps to a given distance would indicate the absence of visual distance perception, but this could not hold under other conditions, for instance when the animal may also rely on learning and kinesthetic cues. However, under both test and control conditions, with random and successive orders of presentation of distances respectively, the greater proportion of jumps were inaccurate. In this respect, therefore, the accuracy criterion fails to be discriminating. But it does show the distinct advantage of the successive order over random order of presentation in the case of overjumps and underjumps and, for this reason, it would appear to be a valuable supplement to the jump force criterion.

12. It is suggested that a combined criterion, weighting accuracy and jump force, be used for further experimentation but no method of weighting is offered.

13. Assuming the two criteria of jump force and accuracy are adequate, it is finally concluded that rats cannot perceive distance through their visual apparatus alone to any marked extent, but require learning and kinesthetic cues.

REFERENCES

1. EDWARDS, A. L. *Statistical Analysis*. New York: Rinehart, 1946.
2. LASHLEY, K. S. XI—A preliminary test of innate organization. *J. Genet. Psychol.*, 1934, **45**, 136-144.
3. MAIER, N. R. F., & SCHNEIRLA, T. C. *Principles of Animal Psychology*. New York & London: McGraw-Hill, 1935.
4. MARQUIS, D. L. The criterion of innate behavior. *Psychol. Rev.*, 1930, **37**, 334-349.
5. PETERS, C. C., & VAN VOORHIS, W. R. *Statistical Procedures and Their Mathematical Bases*. New York: McGraw Hill, 1940.
6. RICHARDSON, F. A study of sensory control in the rat. *Psychol. Monog.*, 1909, **12**. Pp. 124.
7. RUSSELL, J. T. Depth discrimination in the rat. *J. Genet. Psychol.*, 1932, **40**, 136-159.
8. SHAAD, D. J., & HELSON, H. Group presentation in the method of constant stimuli as a time-saving device. *Amer. J. Psychol.*, 1931, **43**, 422-433.
9. WASHBURN, M. F. *The Animal Mind*. New York: Macmillan, 1926.

Department of Psychology
State College of Washington
Pullman, Washington

DEVELOPMENT OF PERCEPTION IN THE YOUNG CHILD AS OBSERVED IN RESPONSES TO THE RORSCHACH TEST BLOTS*

Gesell Institute of Child Development

LOUISE B. AMES, JANET LEARNED, RUTH METRAUX, AND
RICHARD WALKER

A. INTRODUCTION

The present study traces the development of perception in the child from 2 to 10 years of age as he responds to the standard ink blots which make up the well-known Rorschach test. This test, though considered by some to be a test of the imagination, is felt by Rorschach himself to be clearly a test of perception. He states (4, pp. 16-18) that there "can be no doubt that this experiment can be called a test of the perception power of the subject."

The interpretation of the chance forms falls in the field of perception and apperception rather than imagination. "Perceptions arise from the fact that sensations, or groups of sensations, ecphorize memory pictures of former groups of sensations within us. This produces in us a complex of memories of sensations, the elements of which, by virtue of their simultaneous occurrence in former experiences, have a particularly fine coherence and are differentiated from other groups of sensations. In perception, therefore, we have three processes: sensation, memory, and association. This identification of a homogeneous group of sensations with previously acquired analogous complexes, together with all their connections, we designate as apperception. It also embraces the narrower term of perception."

If perception can also be called an associative integration of available engrams (memory pictures) with recent complexes of sensations, then the interpretation of chance forms can be called a perception in which the effort of integration is so great that it is realized consciously as an effort. This intrapsychic realization that the complex of sensations and the engrams are not perfectly identical gives the perception the character of an interpretation.

The present investigation has analyzed the reactions of presumably normal New Haven children of upper middle class status and of high average or superior intelligence. Fifty subjects were observed at each six-monthly

*Received in the Editorial Office on May 5, 1951.

age level from 2 to 6 years and at yearly age levels from 6 through 10 years. Complete analysis of the entire Rorschach response of these subjects is reported elsewhere (1). This paper reports, developmentally, i.e., with emphasis on age changes, those aspects of the response which pertain most immediately to the perceptual abilities of the child.

It is planned that a later study will investigate in more detail the perceptual reactions of the child to the Rorschach cards. Such a study would determine specifically just which portions of the cards attract the most response and what sized and what shaped stimuli are most effective in eliciting reactions at the different age levels. An effort will also be made to correlate individual differences in type of visual functions with the type of visual stimuli in the Rorschach to which the child responds.

The present study, however, considers the grosser aspects of the Rorschach response such as: Does the whole card or some part detail determine the response? Does the color or form of the blot influence the child's answers? Is the form accurate or inaccurate? What type of content does the child perceive in these actually unstructured stimulus blots?

It is believed by the writers that such a preliminary report may usefully precede a more detailed analysis of the blots considered purely as perceptual stimuli.

B. AGE CHANGES IN PERCEPTION

It has been observed (3) that the behavior of the child as well as his physical organism develops by means of structured, patterned stages. Higher psychical manifestations as well as grosser physical behaviors appear to conform to this basic rule. This patterned development is determined largely from within. The same stimulus can and does evoke from any given child different responses at different stages in his life cycle.

Thus we find that the standard Rorschach inkblots, when presented year after year to the same child, elicit quite different perceptual responses as the child matures. When they are presented to a group of say 2-year-olds we find rather marked central tendencies in response which change in a somewhat predictable fashion at $2\frac{1}{2}$, 3, $3\frac{1}{2}$ years and thereafter.

These characteristic age changes in response to the Rorschach which we have observed are presented herewith.

1. *Two Years*

The 2-year-old tends to respond to the Rorschach blot as a whole. He names what he sees in a single-word answer. The response at this age often appears to be a general impression rather than an accurate designation.

The card seems to call to the child's mind the general idea of doggie, kitty, tree and the like, rather than specifically resembling the concept.

The child apparently does not perceive in any accurate detail. If asked by Examiner where he sees something he points to some part of the card, apparently at random.

Magic repetition, in which the child calls several cards in series by the same name, whether or not the concept is appropriate, suggests the lack of accuracy and refinement in his verbalized perceptions at this age. It also suggests that the child's tendency to react perseveratively may predominate over actual visual stimuli in determining his perceptions.

Only half the concepts are identified accurately, further suggesting the vagueness of perception.

Virtually none of the 2-year-olds see movement in the cards. Color responses occur to the extent of only .6 per child, and when the child sees color, he very rarely sees form at the same time.

The frequent occurrence of the adjective *broken* suggests at least a vague concept of parts and wholes. The frequent identification of "house" in the tiny central section of Card VII indicates some ability to note and identify tiny sections of a stimulus.

At this age when the child does respond to parts of the card rather than to the total whole he frequently notes twoness. This is most frequently expressed by the conventional plural but is often indicated by such a repetition of the noun as "doggie—doggie," or by the use of a singular noun meaning nevertheless two separate animals.

At this age such answers as "foot snow," "turkey mountain" may be more immaturity of expression than approach to the "contamination" characteristic of later ages.

2. *Two and One-Half Years*

Perception at this age appears to be very similar to that of the 2-year-old. However, though accuracy of form remains only about 50 per cent as at the preceding age, the wider variety of content, the slight diminution of magic repetition, and the slightly larger number of detail responses suggest a less stereotyped and slightly more realistic identification of forms. There are half again as many popular forms as at 2 years, indicating greater similarity from child to child in response.

Color is now seen in connection with form, CF responses now predominating over pure color responses. However, a large number respond to color merely by naming the color (an average of .54 per child).

Plurals are now expressed by the usual plural, or the phrase "doggie, another doggie."

At this age we see the beginning of the perceiving of some one part or several parts of a total whole, i.e., arm or leg of a person, without connecting this part with the whole to which it belongs. This manner of perceiving appears to be contrary to the customary method of perception, but does occur in some children, particularly during the first five years.

Immediate perceptions only rarely recall to the child related events or concepts. An occasional child, on naming the blot, a Christmas tree, for example, may say, "I got a Christmas tree at home," but this is unusual.

3. *Three Years*

Perceptions are still primarily global, though we now have an average of .88 small or unusual details per child. Accuracy is slightly increasing, F+ % or correct form per cent being 60 per cent. The child still identifies his percept by a single word name, though in a few instances he may express a slight qualification by the phrase, "Looks like a ———." Ability to give combinatory wholes (DW) has increased to an average of .46 per child. In a few instances (.38 per child) the white parts of the blots (S) rather than the shaded or colored parts are identified.

The naming of details of a person or object without building up to the total whole occurs here slightly more than at 2½ years. Thus "Cecily's feet. And her hands. Where are her eyes?" may be the entire response to a blot.

Plurals are expressed by the usual plural; the phrase "doggie, doggie," or the phrase "doggie, another doggie."

The ability to note fine shadings within the black and gray portions of the blots first occurs here to any noteworthy degree (an average of .58 times per child), with texture responses the leading type of shading response.

Stereotypedness of response, and with it inaccuracy, is diminishing. Only 68 per cent of the group use magic repetition, and 26 per cent of the subjects are able to give more than three answers without perseveration.

There is now, successful or not, a much greater tendency to think about, weigh, evaluate, and even question responses. Though simple one word naming of the blot still persists, several say, "I think it is ———," "It's supposed to be ———," "Probably a ———." Or they give their answer in the form of a question: "Is it a doggie?" The personal "It looks to me like ———" is beginning to come in.

Twice as many movement responses are given at 2½ years, but such

responses are still very minimal. Color responses are about the same as at the preceding ages, but the fact that FC responses now hold second place among color responses shows that form is coming to play a more important rôle in the definition of color percepts. However, color naming, as at 2½ years, remains high, an average of .40 per child.

Such descriptions as "a furniture leg house" suggest the beginning of later confused perceptions.

4. *Three and One-Half Years*

Perceptions continue to be global,—in fact by a larger margin than ever before. W% (or per cent of wholes) is the highest to date, and D% (per cent of large details) the lowest. The child not only tends to perceive wholes rather than details, but the ability to produce combinatory wholes has increased sharply. The drive to combine parts into wholes is, however, often considerably greater than the ability to do so accurately. This results in strange verbalizations which presumably reflect perceptions which are probably quite different from "normal" adult perceptions. The 3½-year-old is known to have difficulty in correlating his visual patterns, and this is reflected clearly in the Rorschach response. Thus we obtain for Card III such a response as:

"Butterfly. What are dese? The eyes (center dark). Somebody seeing a butterfly. Hands (top red). These must be a feet be a hands. These must be hands (the usual hands). Funny seeing hands be a feet."

A further example of confused perceptions (and conceptions) is that the child of this age frequently confuses the animate with the inanimate. Inanimate objects are endowed with human appendages: (IV) "A fireplace. Feets dat hold it up. Hands of the fireplace."

Furthermore the child seems to feel quite free, in describing a blot, to describe things not actually seen in the blot:

V. "A big horsie. Horsie's neck and foots and then de horse doesn't have them in now; went home to get his supper and came back and went to soldiers."

So-called Do responses (oligophrenic details) occur here conspicuously, on an average of .36 per child. The child sees and names parts of a customary whole and may or may not actually see or name the total whole:

III. "That looks like a duck. That's a ribbon. That's a pointed nose (man's face). Feet. There's another feet (center black). They're playing with the ribbon. Ribbons (top red). (Asked by Ex. who is playing with the ribbon, replies, 'Men with sharp noses.')

At this age the child fails to differentiate what he sees in pictures and what he sees in real life. Thus he comments of the bears on Card VIII, "Ouch! They bit me," or on Card I, "Looks like a face. It *is* a face."

We should, however, recognize all these confusions and inaccuracies of perception as a step forward in maturity of response over the simpler and more "accurate" response of the 3-year-old.

An extreme aberration of perception occurs when the 3½-year-old in this situation, as in others, complains with no *apparent* justification, "I can't see."

Independence of the usual visual spatial orientation is suggested by the fact that children are beginning to see and identify forms which are upside down, without commenting on the fact of the upside-downness.

Shading responses occur here to the extent of .36 per child. Diffusion is the outstanding type of shading response, suggestive of the vagueness of the child's perceptions at this age.

However, for all this confusion, perceptions are becoming more accurate. F+ % has increased to an average of 67 per cent. Magic repetition, which implies a predominance of perseverative tendencies over accurate perceptions, has now decreased to the point where only about half of our subjects (52 per cent) exhibit this behavior.

More children now see movement in the cards than formerly, animal movement responses now occurring on an average of .70 per child, more than twice the amount at 3 years. Color responses are by far the highest to date—sC being on the average 1.07 per child. CF responses are the predominant type.

The majority of subjects identify the concepts simply by naming the object seen. Plurality is now almost entirely expressed by the customary grammatically correct plural, though some are beginning to use the phrase "Two doggies."

5. Four Years

Manner of perception changes somewhat at this age level. The child exhibits much less of a tendency to see wholes, and gives much more attention to details, especially to tiny details. Whole responses drop from 60 per cent to 52 per cent of the total. Large details rise from 34 per cent to 38 per cent, and tiny details from 6 per cent to 10 per cent.

Perceptions tend to be accurate and very specific. "A penguin. *With no doubt at all.*" However, "confabulation" is frequently present. The 4-year-old is a creature of often rather unbridled imagination. To what extent his "confabulations" are matters of perception and to what extent

of imagination is difficult to determine, for instance in such a "confabulation" as the following:

VII. "Looks like a stone, these are two stones (tier 3) and dust (tiers 1 and 2), a coming up to stones. Little boy moving inside stone but you can't see him cause stones are hard to open. He's got a little bed and table inside the stone and the dust hears him moving. Moving about playing with his toys."

As at 3½ years, distinctions between animate and inanimate are not entirely clear cut. Plants and objects continue to be perceived, or at least described, as having human appendages.

Efforts to achieve greater clarity of perception or conception are frequently expressed in the child's verbalization:

VII. "That looks like a bunny's ear and tail and nose, and tails. What he standing on? Rocks. That's the only way to stand up or you could stand on a ladder. Not bunnies, just grownups."

II. "That looks like a kitty cat. Never heard of a kitty cat having red ears. I had a kitty cat but he moved away."

Four years is a high point for perceiving serially parts of a total whole which may build up to the total whole or may simply be left as a series of parts. Such responses are scored *Do* but probably do not have the same significance as does the *Do* response in adults. Thus on IV: "Shoes. Feet. Hand. Faces. Neck. No mouth." (Whose?) "His face."

Though position alone seems to determine many responses at this age, i.e., portion at the top of a card being called the head,—much of the perception appears to be independent of the usual spatial orientation. Several children now see things upside down without any apparent need or wish to turn the cards to obtain correct orientation. Thus one child, holding Card IX in the usual orientation, describes the blot as "Dat looks like two babies looking down from a tower," the pink portion of the cards, at the bottom, being the babies, and the green portion the tower.

The 4-year-old child tends toward symmetry of posture, movement, and perception. In the Rorschach he tends to give counterbalanced answers. Thus he names a part of the blot "doggie" and its counterpart on the other side of the blot "kitty." Or, "This is a man and this is a girl." Plurality is expressed chiefly by the usual grammatically correct plural. However, this is the first age when we find conspicuous use of the counterbalanced plural.

The high animal per cent (59%A), indicative of marked stereotypy, and predictability, of response, suggests a certain uniformity of perception

from child to child, as does the fact that 15 per cent of any child's answers, on the average, are popular to the extent that they are given by one child in three.

Only 40 per cent of the subjects now persevere.

Movement responses have nearly doubled since $3\frac{1}{2}$ years. Color is seen to about the same extent as formerly.

The majority of children still simply name the blot. However, at this age for the first time comparing the blot to the concept with such phrases as "Looks like" is the second most prominent type of identification.

6. *Four and One-Half Years*

At this age there is a strong tendency toward the perception of wholes: 56 per cent of the responses given are global and only 33 per cent are large details. However, the percentage of tiny detail responses remains high—11 per cent. Thus though global perceptions predominate, some children select a tiny portion of the blot, name it and ignore the rest of the card. However, Do responses, i.e., naming small separate parts which may or may not add up to a total whole, have decreased markedly.

Perceptions at this age do not appear to be clearcut. Whether this uncertainty is actually a quality of the perception or is primarily an emotional factor is not definite, but it may well be at least related to perception. Many children, for example, give an initial refusal even though they are actually able to give a good response. Others criticize their response having given it, while others give a response and then ask confirmation of its accuracy from the Examiner. Some in their uncertainty tell about things they *would* be able to identify. Thus:

V. "Uh! I know what these are but I can't think of the name. If you had a house I could tell that easily because if it had a chimney and smoke I could tell it's a house."

Some seem to be trying to confirm the accuracy of their perceptions by assuring themselves and the Examiner that the real object looks like what they perceive in the cards:

VI. "A tree. Because I could tell because trees have this kind of long thing to hold the branches on, like a great big stick."

Not only is the perception of the $4\frac{1}{2}$ -year-old uncertain, but it appears also to be extremely *confused*. Some of this confusion appears to be caused by the fact that the child is striving to give a whole response including white portions of the test (as cat's face for Cards I or III), but still and

at the same time sees the side figures separately as human or animal figures. Others see "animal with two noses," "fish with two heads," "bug with two tails." Other responses approach the full "contamination" of 5 years. Thus we get: "A cabbage—nose, eyes, cheeks," "A bridge—its legs." "Confabulation" occurs, though less than at 4 years. Other perceptions typical of this age level, neither "contaminated" nor "confabulated," can best be classified as merely *strange*:

VIII. "Here's a butterfly flying up (bottom pink and orange) and he has to keep this thing right on a straight line (center). He can't put anything in that, it's a snap thing, you can put a paper there to hold it."

As earlier, the child can perceive and identify things which are in an upside down orientation, but now he mentions the fact that they are upside down.

There is a strong tendency here to see objects in the white rather than in the colored or black and gray portions of the blots.

Perseveration is decreasing—only 30 per cent of the subjects persevere; and stereotypy is low, A% having decreased in six months from 56 per cent to 44 per cent.

There is somewhat less movement than at 4 years but much more response to color—that is, perceptions are much more determined here by external than by internal factors. Color responses are frequent and largely unmodulated—CF responses predominating, C responses coming next, and FC occurring least of any.

Shading responses occur to the extent of .58 per child. Diffusion and texture responses are the chief kinds observed.

Reversal of the usual figure and ground and naming the white rather than the shaded or colored parts of the blot occurs extensively at this age, both naming of small white parts, or the total WS responses.

Simply naming the blot still predominates as a method of identification, but comparing the blot to the concept by the use of such phrases as "Looks like" is increasing.

In expressing plurality, the usual plural still predominates but the phrase "Two doggies" now occurs more than half as much as the usual plural. This is the high point for counter-balanced plurals, i.e., naming two similar balanced portions of the card "Doggie, kitty," or "Mummy, daddy."

7. Five Years

Type of perceptivity remains primarily global; barely half the group give any small details and 58 per cent of responses are global. The child at-

tempts to encompass the whole of the blot in his interpretation. He responds focally, often picking out an area at the center of the card, then relating the sides to it, with bilateral balance. When he gives a detail response he often chooses details in the center of the card; or if he points out a detail on one side of the card he balances it by mentioning its duplicate on the other side. Sometimes he names these two parts alike. Sometimes he counter-balances his answers: "Here's the mother, here's the daddy," or "That looks like the sun and that looks like the moon." Otherwise, as to plurals, the usual plural and "Two ——" occur about equally.

Accuracy of perception is increasing, the average $F + \%$ being now 78 per cent. In his wish for complete accuracy the 5-year-old often hesitates to give his answer and temporizes with "I don't know" before answering. Or he protects himself with the phrase "It looks *to me* like." Or he may refuse completely. Qualifying concepts occur here more than at any other age except 7 years. The child says, "If you cut this off it looks like ——." Or he may ask the Examiner, "What do *you* think it looks like?" "What does it look like to you?" Or he may give an answer and then criticize his own concept: "That is a butterfly. Too long a ear though. Never saw feet on a butterfly." Indecision is also expressed in the phrase, "I think it's a ——."

Perseveration occurs minimally—only 28 per cent of subjects perseverate; and when there is perseveration, the concept is now at least somewhat adapted to the actual appearance of the blot. Dynamic perseveration, that is perseveration with change of detail as it continues, begins here.

Subjects frequently assume that two parts of a blot go together, by pointing to only one side but saying "Two ——." They also show the ability, on Cards II and III, to take the outstanding parts, bridge the white spaces, and make the whole card into a face, front view. Five has a high generalizing ability—DW is the highest of any age to date, and WS (a total whole based on the white part of the blot) is exceeded only at 7 years. Though WS is high, use of small white space parts is low at this age.

Five's strong tendency to give only one answer to any one picture frequently leads him to combine elements of two separate responses in one response, with resulting "contamination": "Dog-map," "Horse-fly." Or, the response may only approach "contamination," by giving human appendages to plants or objects: "A bridge, its feet"; "A cabbage, its eyes."

Approximately two-thirds of responses at this age are based on form alone. Movement responses though still few, are more than at the preceding age. Color responses are few, there being fewer of every kind of color response

than at $4\frac{1}{2}$ years. Much dislike of blackness of cards is mentioned. Color naming is virtually non-existent. Shading responses occur only on an average of .44 per child, less than at any succeeding age.

Stereotypy is again low (A% 44 per cent), there being considerable variability in response from child to child.

8. *Five and One-Half Years*

Perceptions continue to be primarily global, though at this age there is a good deal of response to tiny details (an average of 12 per cent Dd per child). We note contrast between the perception of total wholes and very tiny details in such contrasting responses as "Eyelashes for the mouth" or "Clam's tongues," and "The world."

Generalizing ability is high with the largest number of DW responses to date. Wholes tend to be broken up and recombined. For instance on Card VI instead of seeing a whole animal the child sees a butterfly on a leaf or a kitten on a rock. The drive to combinatory activity introduces some confusion into the response. "Contamination" is even stronger than at 5 years: "butterfly-map" or "leaf-fish."

The child of this age is rigid, exact, and unmodulated in his perceptions. To him the blot seems to be less of a picture and more of a puzzle. He gives his answer and then asks, "Is that right?" He not only seeks accuracy but achieves it—F+ % is 84 per cent, the highest to date. Semantically he expresses this rigidity and exactness of perception by simply naming the blot. At this age, such qualifying phrases as "Looks like" or "Looks something like" are conspicuously infrequent.

Though some subjects, when questioned, can explain their perception, the majority when asked "How could you tell?" or "What reminded you?" answer simply, "Because," or "Because I can tell by looking at it."

There is much emphasis here on broken or crooked things, or on things with holes in them. As one investigator has described it, the cards almost appear to shatter before the child's very eyes.

Response to the small white areas of the blot, rather than to the black, gray, or colored areas, reaches an all-time high with an average of .60 S per child.

Twenty-six per cent of subjects still persevere, but the perseveration is most frequently dynamic, i.e., changing slightly from blot to blot.

There are slightly fewer human movement responses than at 5 years, slightly more animal movement, many more inanimate objects in movement. However, there are many more color than movement responses, and many

more color responses than at the preceding age,—sC is 2.26, the highest of any age to date. CF responses definitely predominate; C come next in frequency. There are few Fc responses. Shading responses occur more than at any other age to date—on an average of .60 per child.

There is marked variability of perception at this age. A% is only 41 per cent, lower than at any other age in the first 10 years. There is also great variability in any one record, from "good," plain, popular forms to vague unclear responses.

As to plurals—for the first time the phrase "Two ——" predominates over the grammatically correct plural.

9. *Six Years*

Over half the 6-year-old's perceptions remain global, though at this age response to tiny details reaches an all-time high—15 per cent. Tendency to produce combinatory wholes is high—an average of 1.42 per child, though lower than at surrounding ages. Giving part responses where the normal adult sees a whole is very high, an average of .44 Do per child.

The 6-year-old is matter of fact, factual and realistic. His perceptions tend to be accurate—F+% is 81 per cent. Not only that, but he makes much use of the word "real." He sees "real birds," "a real baby." He likes the card that looks "realest." Conversely, there is a marked absence of the "contaminated" and "confabulated" responses given earlier. There are only three "confabulated" responses here and no "contaminations."

Six also appears to be sure of his perceptions. Refusals and denials are both low and there are almost no qualifying phrases used. The child identifies blots either by simply naming the blot or by using the phrase "Looks like."

Perceptions are not only accurate but predictable and somewhat stereotyped at this age. A% (48 per cent) is higher than at any other age except 2 and 4 years. P% is 28 per cent. Responses are perhaps more standard and predictable at this age than at any other. On Card I, for example, 36 out of 50 children give responses falling in only three categories: (a) things with holes in them; (b) butterfly or bird; (c) face or mask.

Only 18 per cent of subjects persevere, and perseveration when it occurs is static, not dynamic as at preceding and following ages.

There is a strong tendency toward reversal of the usual figure and ground. Identifying the white portions of the blot instead of shaded areas (S) occurs here more than at any other age except $5\frac{1}{2}$; and WS responses occur more than at any ages except $5\frac{1}{2}$ and 7 years.

At this age, only 60 per cent of responses are determined by form alone, there being both more movement and more color responses than at any other age to date. As to color, CF responses lead, FC comes second. Shading responses occur more than at any other age to date—an average of .68 per child.

10. *Seven Years*

Perception still remains primarily global, but the perception of large details has now increased and is higher than at any age since $2\frac{1}{2}$ years. Tiny details are noted much less frequently than at the preceding age, having dropped in frequency of occurrence from 15 per cent to 8 per cent. The drive to generalize and to produce combinatory wholes is at an all-time high with an average of 1.98 DW per child. WS also reaches an all-time high of 1.02 per child. The 7-year-old can handle several discrete details at the same time, building them into a whole response, or combining them below the global level. With his ability for quick organization he often cuts across the usual boundaries of a blot to organize it into a whole form. On Cards II and III for instance where many subjects see two human or animal forms, Seven pulls the sides together, fills in the white spaces and produces a single face.

Not only are WS responses frequent, but we have here for the first time a complete reversal of figure and ground, white spaces not only being seen as holes or caverns but themselves becoming actual objects such as tops, ducks, bellows.

The drive toward generalization, as at several earlier ages, often exceeds the ability to generalize accurately, and results in a large amount of "confabulation," much more than at 6 years. For all this "confabulation," accuracy of response is high, $F+\%$ being 82 per cent.

Perseveration occurs in more than half the cases at this age and in most instances it is an extremely dynamic perseveration, the basic concept remaining fairly consistent but its attributes changing.

Qualification of own perceptions occurs more than at any other age to date. The child qualifies his response with such phrases as: "Might be," "Could be," "Looks something like," or "If you cut this off it would look like ———." The 7-year-old is highly critical of the differences between the actual shape of the blot and his concept, and he tries to reconcile the differences by verbally adding to or subtracting from the blot. Or in his uncertainty, the child gives his response, then denies or changes his answer. Simply naming the blot remains the chief method of identification, though

comparing blot to concept by use of such phrases as "Looks like" comes a close second.

Though individual records are not stereotyped—A% is very low for this age range, only 42 per cent—the number of popular responses is higher than at any other age to date—27 per cent. Not only the actual things seen may be the same from one child to another, but even the entire phrasing used to describe the concept.

Form alone determines only 52 per cent of the responses, an all-time low. Both color and movement responses of all types (except CF responses) occur more here than at any other age to date. Color responses are frequently combined with inanimate movement—thus blood and paint are splashed, fire burns. Shading responses reach an all-time high, differentiation of details within the blot, and texture responses, leading.

There is at this age a strong expression of the morbid. Decay, damage, and mutilation are all conspicuous themes. The following are quite characteristic perceptions: "Dinosaurs pulling up blood skeleton," "Decaying bones," "It looks as though the bones had started to deteriorate—like rot over here."

Unlike Six, many Sevens are able to give a clear-cut answer when asked where they see a certain thing or how they know it was such and such a thing.

11. *Eight Years*

Global perception once again leads by a large margin (55%W), and both large and small details decrease in number. There is somewhat less drive toward the production of combinatory wholes (1.50 per child), but such wholes when produced are usually of reasonably good form. "Contamination" and "confabulation" no longer characteristically occur.

The 8-year-old seems to feel a great need for accuracy of perception and response, and to a large extent he achieves this accuracy. F+% is 87 per cent. In his search for complete accuracy, the child expresses much uncertainty and qualification. This is the highest point to date for qualified responses—"Looks something like," "Might be," "Could be," "I guess," and the first age at which comparing the blot to the concept by use of such phrases as "Looks like" predominates over simply naming the blot, as earlier.

Temporizing is frequent, followed by a good answer, but complete refusals and denials of own response are also high.

Eight is the high point for direct suggestions for changing the blot to make it fit the child's own concept. He suggests, "If you cut this off," "If this was different here and this part wasn't in it."

In spite of uncertainty and temporizing, responses are more clearcut and less confused than at either 7 or 9 years.

Eight like Four seems to feel a need for exactness, and often appears to feel, in spite of instructions to the contrary, that the blots actually represent some one real thing. He responds and then asks, "Did I guess right?"

The blots appear to engender unpleasant feelings in many 8-year-olds. To what extent this is a matter of perception is not certain. Some say directly of the cards, "Oh, I hate these!" Others describe the cards as "horrible" or, as at seven, see morbid and gruesome details: "Man's legs bleeding and the bones of his feet," or "Hog's body cut out as a rug, blood coming out." There is a marked increase in anatomical responses.

Form determines a high percentage (58 per cent) of the answers. Movement responses are somewhat fewer than at 7 years, and color responses are much fewer. However, there are relatively more FC at this age than at preceding ages. Shading responses are frequent, nearly one per child.

Perseveration is no longer a conspicuous factor. Disappearance of perseveration, "contamination," and "confabulation," along with an apparently strong drive to give accurate responses, and a diminution in morbid and gruesome responses result in a much more clearcut response at eight than at seven years of age.

12. *Nine Years*

At nine years there is a definite shift in type of perception. Here for the first time, as in adult perception of the Rorschach, the subject sees more details than wholes: 48 per cent of forms seen are now large details, only 42 per cent are wholes. However, the drive toward production of combinatory wholes is high, there being an average of 1.64 DW responses per child. These, like other forms, are described with considerable accuracy, $F+\%$ being at this age 84 per cent.

The outstanding characteristic of perception is the child's uncertainty about what he sees, which is even more marked than at 8 years of age. In fact nine years is the high point for the whole age range for qualifying responses. The child says, "It looks something like," "It could be," "It might be." Suggestions for changing the blot to fit the child's concept are very strong here: "Cat's face without these. Without these knobs up there. And make it circular sort of like in here and cut this off."

Though $F+\%$ is high at this age, perceptions often appear to be considerably confused. The child frequently denies his own response, and responses are often difficult to score because it is hard to tell what the

child means. "Oh, this is a Viking's head; no I don't think it's a Viking's head. Map of Wonderland because I can't think of anything else."

There is considerable attention paid to the position of cards, and there are many comments as to whether cards are upside down or right side up, or statements to the effect that "If you turn it like this then it looks like . . ."

Many of this age confirm their perceptions by referring to outside sources: "I saw a picture like that in a book," or "I know because I saw some in Peabody Museum."

Perceptions are variable at this age and individual differences are marked. Perseveration does not occur. As to content of perception, anatomical and map responses both reach high points here.

Form alone determines a larger number of perceptions than at any age since 5 years, 67 per cent of responses being determined by form alone. Movement and color responses, though relatively fewer, have increased in actual number since 8 years. Color responses are particularly unmodulated, FC responses being exceeded by both CF and C responses. Shading occurs conspicuously (.84 responses per child) though less than at the two preceding ages.

13. *Ten Years*

Once again perception is primarily global, 52 per cent of responses now including the total blot, 40 per cent being large details, and 8 per cent tiny details. Drive toward production of combinatory wholes is low, lower than at any age since 5 years. Reversal of the usual figure and ground is relatively low (an average of only .32 S per child); WS is moderately high—.82 per child.

At this age for the first time, movement responses predominate over color responses. Both animal and human movement responses have increased since the preceding age. Response to color stimuli is sharply reduced, sC being lower than at any age since 4 years. CF responses still predominate; FC responses come second and are numerous. Shading responses are fewer than at any other age since 5½ years, texture responses for the first time constituting the main type of shading.

Suggestions for changing the blot to fit the child's concept, characteristic of the immediately preceding ages, are now sharply reduced. The actual visual stimulus now seems to carry more weight, and if concept and blot do not agree, the child gives his response and then changes his concept to fit what he perceives, rather than changing the blot.

There is decreased productivity at this age, and much more similarity

in response than formerly from child to child. Number of banalities is higher than at the preceding ages and different content categories used are fewer.

Ten shows an increased critical ability. $F+\%$ is now 89 per cent, the highest to date. The high number of refusals, .60 per child and the most at any age since 3 years, now probably indicated less unwillingness and inability to respond than a wish to be clear and accurate. Greater critical ability is also suggested by the considerable attention paid to the details of facial expression.

In spite of the greater demand for accuracy, general uncertainty of response continues to be characteristic. The child gives a response, questions it, returns to his first idea, tempers his statement with such a phrase as "sort of." Comparing blot to concepts by such phrases as "Looks like" is the chief method of identification. This is one of the three leading ages for use of the phrase "Looks something like."

A rather unusual characteristic of perception at this age is the child's frequent confusion of human and animal figures. He also tends to mention things not actually seen in the pictures. "Hunter opening up an animal. The hunter isn't there."

There is a strong emphasis on twoness, and also on things being attached or pulling apart. . . . Anatomical and supernatural responses are both characteristically frequent.

Orientation is a matter of considerable concern. Card turning occurs more than at any other age to date and even when subject holds the cards rightside up he may describe the blot as "Two upside down boys." Size also appears to be of considerable importance and is frequently mentioned: "Sort of oversized crabs."

C. SUMMARY

The perception of the young child, as observed in his responses to the Rorschach test blots, are primarily global. That is, the child is more influenced by the total configuration of a visual stimulus than by any specific details. At every age from 2 through 10 years, with the sole exception of 9 years of age, 50 per cent or more of the child's responses encompass the entire blot. Responses to what are commonly described as "large details" of the blots range in number from 33 per cent to 44 per cent except at 9 years when they rise to 48 per cent of total responses. Responses to tiny details are fewer, ranging with increasing age from 4 per cent to 15 per cent.

Whole responses which are made up of separate parts organized more or less successively as wholes, occur minimally in the first three years; but from $3\frac{1}{2}$ to 10 years occur on the average of about one per child per record, surpassing that number especially at 7 years when they occur on the average of two per child. Whole responses in which the white area of the blot plays a primary part do not occur appreciably until $4\frac{1}{2}$ years when they occur to the extent of .62 per child, increasing thereafter to a high point of 1.02 at 7 years and decreasing slightly to .82 at 10 years.

Responses to single white areas of the blot (that is reversal of the usual figure and ground) occur less than half a response per child except at $5\frac{1}{2}$ years when they rise to .60 per child. Do responses (responses to only part of a body where usually a whole body is seen—that is seeing an arm or leg where most observers see a total figure) occur to the extent of more than half a response to a child only at 9 years. In the adult, these responses are considered to indicate suppression of the unnamed remainder of the blot. In children, especially at $2\frac{1}{2}$ through 5 years, they may indicate that the child is perceiving serially parts of the whole which he may or may not succeed in building up to a total whole.

Form alone (as opposed to movement, color, and shading) determines a decreasing percentage of responses as the child grows older. At 2 years of age, 90 per cent of responses are determined by form alone. This percentage decreases gradually to a low point of 52 per cent at 7 years, then rises to 63 per cent at 10 years. Correct form ($F+\%$), conversely, rises steadily with increasing age from 54 per cent at 2 years to 89% at 10 years. That is, as might be expected, the child's perceptions become increasingly accurate as judged from an adult viewpoint, as he matures.

Responses to colored parts of the stimulus increase fairly steadily from 2 years to a high point at 7 years, then decrease until 10 years. At 2 years and *only* at 2 years, pure color responses predominate. At *all* other ages, CF responses, i.e., responses in which color predominates but some form is present, are the leading type of color response. Pure color responses hold second place at $2\frac{1}{2}$, $3\frac{1}{2}$, $4\frac{1}{2}$, $5\frac{1}{2}$, 7, and 9 years. At other ages FC responses, i.e., responses which contain color but in which form predominates, come second.

Responses to the colored parts of the plates are generally considered as direct responses to environmental stimuli, and most Rorschach workers consider them as emotional stimuli. Color responses are considered as a "balance between the strength of the emotion excited and the mental control with which it is met. . . . A pure C answer is considered to be a direct,

uninhibited reaction to the stimulus with no moderating or controlling factor operating. . . . In CF responses the emotions are in command but some element of control is present. . . . In FC responses, emotions are controlled and subordinated to consideration of the mores" (2, pp. 43, 44). The degree of intellectual control over the emotional responses appears, from our figures, to increase with increasing maturity.

Movement responses, in contrast to color responses, have their source not in the purely external stimulus but in the "inner" life of the subject. Identification and inner creativity are considered to be their major components, and the individual who gives movement responses is thought to be employing creative imagination (2, p. 39).

We have found human movement responses to be virtually absent up till 6 years, occurring no more than to the extent of $\frac{1}{2}$ a response per child. At 6 years there is an average of one per child, and by 10 years, $1\frac{1}{2}$ per child. During the first five years, animal movement occurs about twice as much as does human movement. From 6 to 10 years, this ratio decreases steadily till at 10 years we observe 1.66 instances of human movement per child and 1.74 instances of animal movement. Instances of inanimate objects in motion occur less than once per child at any age, the range being from .05 per child at 2 years to .36 per child at 10 years. The high point is .82 at 7 years.

Shading responses, that is responses defined or differentiated by shading values of the black and gray areas of the blots range from a frequency of .24 at 2 years to .60 at ten. Their high point is at 7 and 8 years when they occur to the extent of 1.14 and .92 per child respectively.

As to outstanding types of content projected onto the blots by children,—animals are at all ages outstanding content category. A% is 50 per cent at 2 years and decreases only slightly with increasing age, being 49 per cent at 10 years. Since the cards lend themselves most readily to animal interpretation, animal responses tend at all ages to be the most frequently given type of response. Since the giving of such a response requires little or no originality, a preponderance of such responses suggests a certain obviousness and stereotypy of thinking.

Of other leading content categories, *objects* at all ages is one of the four leading types. *Plants* are conspicuous through $5\frac{1}{2}$ years; and at some ages *nature* responses are prevalent. *Humans* come in as one of the four main categories at 3 years and thereafter; and at 8, 9, and 10 years, *anatomy* responses are also strong.

The earliest method of identifying the blot consists of simply naming

it. Next comes identification by stating that the blot is something. Then comes comparing blot to concept by use of the phrase "Looks like." Qualifying concepts which indicate that the child is not certain of the accuracy of his perception occur most at 5, 7, 8, and 9 years of age.

Much of the child's perceptivity, as measured by the Rorschach test in the first 10 years of life, appears to be influenced by strong, developmentally determined reaction tendencies rather than solely by the actual stimulus value of the cards themselves. Thus perceptions in the first three years tend to be vague and inaccurate. The child identifies the whole blot, regardless of its actual appearance, as *doggy*, *kitty*, *tree*, or *birdie*. If the 2-year-old is asked where he sees the object named, he points at random. Efforts to identify and describe, and particularly efforts to combine all parts of the blot into a whole, often exceed ability.

By 3 years, however, many can point out accurately the portion or portions of the blot which they are identifying. The general inaccuracy of early perceptions is suggested by the large amount of "magic repetition"—perseveratively calling several cards in a series by the same name whether appropriate or not—which occurs in the first four years. This predominance of magic repetition suggests that the child's tendency to respond to any situation perseveratively predominates over the actual visual stimulus in determining his reported perceptual responses. Though the amount of magic repetition decreases steadily, even as late as 4 years it occurs in 40 per cent of subjects. Thereafter it drops to 30 per cent at $4\frac{1}{2}$ years, 28 per cent at 5 years. From this age on, especially at $5\frac{1}{2}$ and 7 years, when perseveration does occur it takes the form of "dynamic perseveration." That is, though the same basic concept is repeated, it changes somewhat from card to card, influenced by the actual appearance of the blot.

Three-and-a-half introduces into the blot, and describes, things which are not actually present in the blot. He also appears to confuse pictures of objects with the actual objects: "Bears. Ouch! They bit me!" The child of this age often finds it hard to correlate his part-perceptions into a clearcut whole. He also confuses the animate with the inanimate, endowing inanimate objects with human appendages. At this age, too, the child's perception often appears to be largely independent of the usual visual-spatial orientational cues. Many children identify objects which are obviously seen in an upside-down orientation without apparently feeling any need to turn the card or to mention the unusual orientation. This suggests the possibility that the child may in the beginning see objects discretely and apart

from what we consider to be their "correct" orientation, and that "correct" orientation may be to some extent a learned response.

The 4-year-old continues to endow inanimate forms with human appendages, and also shows a strong tendency toward "confabulation"—which may reflect confused perception, unbridled imagination, or both.

At 4½, perceptions are uncertain,—the child does not seem at all sure of what he does see—and also confused. Some children see animals with two heads or two tails. Others approach the full "contamination" of 5 to 5½ years in which the child combines elements of two separate responses in a single response, i.e., "Dog-map," "horse-fly," "leaf-fish." Still others give "confabulated" responses, or other *strange* types of response in which it is difficult to tell just what is seen or meant.

The tendency to "confabulation" on the part of the 4 and 4½-year-old leads to descriptions of objects and activities far beyond what would be suggested to the average adult by the outlines of the blots.

At 7 years, the child's interest in the morbid colors his perceptions. At 7, 8, and 9 years he attempts to fit the stimulus to his perceptions by suggesting changes in the shape of the blot: "If you cut this off, it would be ———." At 8, 9, and 10 years of age the child's interest in anatomy influences him to see many anatomical details.

• Thus we find that throughout the first 10 years of life the child's response to the Rorschach inkblots as perceptual stimuli changes in a patterned manner from age to age, approaching though not approximating, the expected response of the "normal" adult.

There is a slight trend with age in the direction of more attention to major details of the stimulus, a strong trend toward more accurate identification of these parts, more response to color aspects of the stimulus, and an increased ability to combine form and color. The tendency to see movement in non-moving stimuli increases, as does the differentiation of fine shadings within the black and gray areas, and the identification of tiny details. Not only is form seen with increasing accuracy but such inaccuracies of perception as perseveration, "contamination," and "confabulation"—in which inner drives appear to cause the child to give responses which frequently have little relation to the given visual stimulus—become increasingly infrequent. Furthermore with increasing maturity the child becomes more certain to see and identify visual stimuli in the orientation which is considered "correct" from the adult point of view.

REFERENCES

1. AMES, L. B., LEARNED, J., METRAUX, R., & WALKER, R. Development Trends in Child Rorschach Responses. New York: Hoeber, 1952.
2. BOCHNER, R., & HALPERN, F. The Clinical Application of the Rorschach Test. New York: Grune & Stratton, 1945.
3. GESELL, A., & ILG, F. L. (In collaboration with L. B. Ames and G. Bullis.) The Child from Five to Ten. New York: Harper, 1946.
4. RORSCHACH, H. Psychodiagnostics. Berne, Switzerland: Huber, 1942.

Gesell Institute of Child Development
310 Prospect Street
New Haven 11, Connecticut

AGGRESSION IN DOLL-PLAY: INTERVIEWS OF TWO- TO SIX-YEAR-OLD WHITE MALES*¹

Department of Psychology, University of Louisville

CAROL H. AMMONS AND ROBERT B. AMMONS

A. INTRODUCTION

Free doll-play has become an increasingly popular projective technique among psychologists interested in studying the behavior of preschool children. Typically, these investigators (3, 7, 10) report that the method is flexible enough to allow free expression of the child's real feelings, and yet at the same time is adaptable to experimental purposes. This is also the conclusion reached in the series of studies of methodological factors which was conducted at the Iowa Child Welfare Research Station under the direction of Sears. In a recent article (12) he summarized their findings and emphasized the usefulness of the technique in diagnosis and research.

Much useful and interesting information concerning children's aggressive behavior has come from doll-play studies. Bach (3) reported that normal children produced "intensely aggressive fantasies." Baruch (4) described physical and verbal aggression toward the family in young children. The amount and latency of expression of aggression have been related to a number of methodological factors, such as the kind and organization of materials, experimenter-child interaction, and length of session (12).

Information about aggression obtained from the free doll-play situation is corroborated by the findings in a number of direct observational studies. Bridges (6), Green (9), and Jersild and Markey (11) found that with increase in age the frequency of crying and physical aggression decreased, while the frequency of verbal aggression increased. Bridges, and Jersild and Markey, also reported an increase in "repression" of aggression with an increase in age. Sex differences in frequency, amount, and form of aggression were noted. Fite (8) pointed out a number of variables which are related to children's attitudes toward aggression. In a general study of 83 patients (3-15 yr.) who were on a children's ward, Bender and Schilder (5) found that the younger children expressed aggression in play and

*Received in the Editorial Office on May 14, 1951.

¹Paper presented in part at the meeting of the Southern Society for Philosophy and Psychology, April 7, 1950.

picture descriptions more freely than did the older children. However, it was possible to elicit aggression from the latter in an indirect manner.

A modification of the doll-play method has been used by Conn (7) and Weiss-Frankl (13) with nursery school children, by Levy (10) in studying sibling rivalry, by Ammons and Ammons (1) in the investigation of parent preferences of preschool children, and by Ammons (2) in an analysis of young children's reactions to differences in skin color and facial features. In this method, the doll-play interview, the doll is interviewed in a miniature toy situation and the child answers for the doll. As this technique has been valuable in studying these areas of child behavior, its application should lead to interesting information concerning aggression.

B. PROBLEM

The doll-play interview method is used to investigate age changes in forms of aggressive responses in preschool-age boys. There is also an evaluation of the effectiveness of this technique in research as compared with direct observation.

C. EXPERIMENTAL ARRANGEMENTS

1. *Materials*

Materials consisted of two dolls and a miniature playground. The dolls, one Negro and one white, were made of pipe stem cleaners, and stood 5 in. tall. They were dressed in denim overalls and cotton shirts. Facial features were painted on pieces of light-weight cardboard, which were stapled securely to the bodies as heads. The Negro doll's head was brown, with characteristic facial structure and kinky black hair. His arms and legs were covered with matching brown cloth. The other doll's head was pink, his legs and arms white. Both dolls resembled four-year-olds.

Around the miniature playground, a pine board $\frac{3}{4}$ in. by 11 in. by 23 in., was a "fence" made with $2\frac{1}{2}$ -in. finish nails and string. In one corner was fastened a 4-in. by 4-in. sandbox in which there was some real sand. A swing, nine small wooden blocks, and a small toy "red racer" auto were put in the other three corners. *E* kept red balloons in his pockets until needed.

2. *Subjects*

Ten boys at each age level, two, three, four and five years, were interviewed. Mean ages in years for the groups were 2.5, 3.3, 4.3, and 5.6. These children were interviewed in three public day-care centers and one public school kindergarten. All were naïve to the doll-play interview situa-

tion except four three-year-olds, five four-year-olds and two five-year-olds. No differences were observed between naïve and experienced Ss. No child was interviewed unless he was willing. There were three refusals at *CA* 2, and one each at *CA* 3 and *CA* 4. Five Ss were discarded, three who were not available for a second session and two because of outside interruptions. Ss were of at least a two-generation "native white" background. Thirty children were living with both parents and 10 with mothers only. The overall occupational classification of the heads of the families was somewhat above the national average, but occupational levels were distributed fairly well over the various age groups.

3. Procedure

E was known to the children as tester and friendly visitor. Each child was approached individually with the statement: "I have some things to play with. How would you like to come with me and play with them?" If the child was willing, he was led slowly to the relatively isolated playroom.

After *E* and *S* had seated themselves on the floor, *E* removed the cloth which covered the materials, saying: "See, here are two little boys on a playground. They have all sorts of things to play with. They can swing, play in the sandpile, ride the car, or build with the blocks." Each activity was pointed out, but materials were not touched and the dolls were left leaning against the fence.

After this warming-up period, the following questions were asked as part of a larger series:

1. They both would like to play with the car. They push each other.
(a) What does this one (colored) do? (b) Who gets the car? (c) Why?

2. They both would like to play with the car. They push each other.
(a) What does this one (white) do? (b) Who gets the car? (c) Why?

3. They both would like to play with the swing. They push each other. (a) What does this one (colored) do? (b) Who gets the swing? (c) Why?

4. They both would like to play with the swing. They push each other. (a) What does this one (white) do? (b) Who gets the swing? (c) Why?

5. This (colored) boy hits this (white) boy. (a) What does he (white) do? (b) What does he say?

6. This (white) boy hits this (colored) boy. (a) What does he (colored) do? (b) What does he say?

7. They play with this big balloon for a while, then they drop it

and it breaks. (a) Who broke it? The teacher comes up and hits one of them. (b) Whom does she hit? (c) Why?

8. This little boy (colored) throws sand at this one (white). What does this (white) boy do?

9. This little (white) boy throws sand at this (colored) boy. What does this (colored) boy do?

Half of the Ss were asked in this order in the first session and half with "colored" and "white" reversed, to avoid a systematic order effect.

At the end of the interview S was permitted to play with the dolls freely for five min. before he returned to the group. After an average of 10.0 days a second interview was conducted. The procedure in both sessions was the same except that order of asking about white and colored dolls was reversed. A more detailed discussion of the procedure can be found in a previous article (2).

D. RESULTS

Data from the second session only are reported here, since the recording and scoring reliabilities were high, and the results for the two sessions were nearly identical. Agreement between two persons who scored the records independently was 95 per cent for the first session and 94 per cent for the second, ranging from 82 per cent to 100 per cent on individual items. Scoring agreement was about the same for all four age groups.

1. *Form of Aggressive Response*

Table 1 shows the form of the aggression and the frequency with which it was mentioned at each age level. There is a large incidence of direct physi-

TABLE 1
FORM AND FREQUENCY OF AGGRESSIVE RESPONSES FOR EACH OF FOUR AGE LEVELS

Form of response	Frequency of response*				Total
	CA 2 (N=10)	CA 3 (N=10)	CA 4 (N=10)	CA 5 (N=10)	
Direct physical counter aggression	31	55	35	24	145
Crying	13		13	2	28
Leaving field	1	8	5	16	30
Displacement of aggression		1	1		2
Inhibition of aggression (denial)			1	8	9
Appeals to adult			8	17	25
Verbal attack			4	5	9
Unscorable answer	2	6	11	12	31
No answer	34	11	5	1	51
Answers scored twice			5	5	10

*There were eight possible scorable answers for each child, and a few were scored twice.

cal counter aggression at all age levels; the three-year-olds showing the greatest amount and the five-year-olds the least. Typical answers given were, "he fights"; "he throws sand at him"; and "he makes a snowball out of paper and throws it at him." Crying is frequent at *CA* 2 and *CA* 4, but entirely absent at *CA* 3, and markedly decreased at *CA* 5. A gradual increase in frequency of leaving the field is noted with an increase in age. Common answers scored leaving the field were "he goes home to his mother," "he runs and plays in the sand," and "he gets the swing" in answer to Question 1 about the car. Displacement of aggressive feelings onto objects in the playground such as the car or the sandbox was infrequent. An increase in inhibition (outright denial) of aggression and in appeals to an adult are noted for *CA* 4 and *CA* 5, but neither appears at the lower age levels. It is interesting that all eight of the four-year-old appeals were to the teacher, whereas this was true of only two of the 17 appeals of the five-year-olds. The other 15 appeals at *CA* 5 were to mothers. These five-year-old children were all in the public school kindergarten and, apparently had not established quite as secure and friendly a relationship with their teacher as had the nursery school groups. Verbal attacks were used by four of the *CA* 4 and five of the *CA* 5 children. Usually, the verbal attack was a command or threat, such as, "Quit it!," "Stop kicking sand!," or "I don't like you!"

The number of unscorable answers increased with an increase in age. The large majority of these answers simply stated the outcome of the activity suggested by the questions, as "He plays with the swing" (Questions 3, 4). However, it is perhaps important that all but one of these answers were given to the first four questions which are of a more general nature than are the second four. This suggests that young children comprehend a concrete alternative more readily, and thus are able to answer more adequately questions which involve limited action. No other reliable differences by age levels were obtained in frequency of a form of aggressive response.

The frequency of refusals to answer shows marked decrease with an increase in age. This probably reflects the child's ability to comprehend and use language more easily.

Scoring agreement for forms of aggressive responses was checked in two ways. One of the *Es* scored the responses a second time 11 months after the initial scoring and found agreement on 94 per cent of the responses. Another scorer,² who was naïve with respect to the experimental design, scored

²The authors wish to thank Mrs. Sylvia Post for aiding in the determination of scoring reliability.

the responses independently, after having the scoring rules explained and scoring a series of responses for practice. Agreement between the *E* and the outside scorer was 89 per cent, ranging from 45 to 100 per cent on individual scoring categories.

A further check on the consistency of the trends in the numbers of answers given for each of the forms of aggression as shown in Table 1, suggests itself. In order to determine whether the same trends in forms of aggression are shown with numbers of *Ss* as with numbers of answers, the number of *Ss* within each age level who showed each form was tabulated. Rank order correlations between the total number of responses of a kind given and the number of *Ss* responding in this way at a given age level were calculated for each of the age groups. For *CA* 2, *CA* 3, *CA* 4, and *CA* 5, the obtained correlations were .93, .82, .79, and .94, with a median of .86. These data indicate relatively high agreement between the two methods of scoring incidence of responses.

2. *Average Number of Words Per Response*

The average number of words per response was calculated for each form of aggressive expression. No significant differences between response categories were obtained within age levels. However, a comparison of the means of all responses given at each of the four age levels showed a gradual increase in average number of words per single response with increase in age. The means obtained for *CA* 2, *CA* 3, *CA* 4, and *CA* 5 were 3.20, 3.88, 4.40, and 4.74, respectively.

3. *Outcome of Aggression*

In Questions 1, 2, 3, and 4, *S* was asked who got the car or swing; i.e., who was successful. In order to obtain a rough estimate of the reliability of the obtained differences in successful outcome, a *t* test of significance was used. There was no significant difference in the success of the white and colored dolls. Rather, success in aggression seemed to depend upon which doll *S* was asked to answer for. The doll for whom *S* described actions tended to be successful in his aggression.

4. *Recipients of Aggression*

Analysis of responses to the reciprocal aggression in Questions 1, 2, 3, and 4 and to the direct aggression in Questions 5, 6, 8, and 9 shows no appreciable differences between the two dolls. As noted in a previous study (2), it would seem that aggression can be freely answered by aggression

in this group, regardless of color, but that when blame is to be placed for something as in Question 7, there is a tendency to scapegoat on a racial basis.

5. *Reasons Given for Successful Aggression*

In general, with increased age there is an increase in the variety of reasons given for why one doll gets the car or swing (Questions 1, 2, 3, and 4). Table 2 also shows a shift from "cause" and "don't know" as the most frequent answers to being recipient of the aggression, to "taking turns," and other responses. This seems to reflect an increased understanding of and greater facility in social situations.

TABLE 2
FREQUENCIES AND KINDS OF REASONS GIVEN FOR SUCCESSFUL AGGRESSION IN
QUESTIONS 1, 2, 3, 4

Reason	Frequency of response by age level				Total
	CA 2	CA 3	CA 4	CA 5	
"Cause"	4	15	3	4	26
Don't know		9	9		18
Wants object	1	9	8	4	22
Recipient of aggression		1	12	12	25
Likes him or "his face"				4	4
Take turns		2		8	10
Inhibition of aggression	1			2	3
"He was bad"			1	3	4
"He was closest to it"				1	1
Unscorable	1	2			3
No answer	33	9	1		43

The kinds of reasons given to Question 7, "The teacher comes up and hits one of them. Whom does she hit? Why?" were similar to those given to the first four questions. Crying, "cause," and "don't know" appeared at the lower age levels (CA 2 and CA 3) most frequently; whereas blaming ("he broke it") was the only reason given at CA 4 and CA 5.

E. DISCUSSION

The results of this investigation are supported by the findings of other investigators who have employed doll-play and various other observational techniques in the study of aggression in children. Among the corroborated findings is that young children express a good deal of aggression (3, 4, 8, 10, 13). The forms which the aggression takes change with age (6, 8, 9, 11), and also become more varied with age (8, 11). In other words, there are a greater number of different responses which the five-year-old makes as compared with the responses of the three-year-old. There is also

a greater inhibition of direct expression of aggression in older children (3, 5, 6, 9, 11). The reasons given for particular behavior reflect increased understanding of and conformity to adult standards in the older children (5, 8). An increase in the number of words per response accompanies an increase in age.

Only when blame was to be accorded for something as in the balloon-breaking situation was there any difference in the successful outcome of the aggression attributable to the "race" of the doll. In all other situations each doll was successful equally often. Since the balloon situation was the only one in which an adult was directly involved, the displacement of the aggression might be related to fear of punishment by the adult (8, 9, 11, 12).

Agreement of findings using the doll-play interview with many of the findings of other studies indicates that it is a valuable technique. It has distinct advantages: simplicity, appeal to very young children, suitability for investigating a number of different areas of clinical interest, and the relative ease with which it may be modified to fit the needs of the clinician or experimenter. Previous studies (1, 2, 7, 12, 13) have evaluated the reliability and validity of this procedure, and their conclusions agree with those of the present paper.

F. SUMMARY AND CONCLUSIONS

The purpose of this study was to determine the course of development of aggressive response with increasing age and to evaluate the doll-play interview as a clinical instrument. Ten boys each at chronological ages 2, 3, 4, and 5 were given two doll-play interviews in which they answered questions about the activities of two dolls on a miniature playground.

Responses to the questions were analyzed by age level and into various categories. It was found that (a) level of expression of direct counter aggression was high for all ages, reaching a peak at CA 3 and diminishing somewhat with increased age; (b) forms and frequencies of aggression changed with age; and (c) crying and direct counter aggression gave way to leaving the field, verbal expressions, appeals to an adult and inhibition of the aggressive feelings. Average number of words per response was found to increase with age. Reasons for successful aggression changed with increased age from "cause," "don't know," and "he wants it" to "take turns," "he was bad," and various aggressions indicating that the recipient of aggression should be rewarded. Number of refusals to answer decreased with age.

It was pointed out that these findings agree with those obtained by other investigators who employed time-consuming observational techniques. This

agreement with the findings with other methods, the high scoring reliability, and the great flexibility as to content support the conclusion that the doll-play interview can be useful in the study of children's behavior.

REFERENCES

1. AMMONS, R. B., & AMMONS, H. S. Parent preferences in young children's doll-play interviews. *J. Abn. & Soc. Psychol.*, 1949, **44**, 490-505.
2. AMMONS, R. B. Reactions in a projective doll-play interview of white males two to six years of age to differences in skin color and facial features. *J. Genet. Psychol.*, 1950, **76**, 323-341.
3. BACH, G. R. Young children's play fantasies. *Psychol. Monog.*, 1945, **59**, No. 272.
4. BARUCH, D. W. Aggression during doll play in a preschool. *Amer. J. Orthopsychiat.*, 1941, **11**, 252-260.
5. BENDER, L., & SCHILDER, P. Aggressiveness in children. *Genet. Psychol. Monog.*, 1936, **18**, 410-525.
6. BRIDGES, K. M. B. The Social and Emotional Development of the Preschool Child. London: Routeledge, 1931.
7. CONN, J. H. The child reveals himself through play; the method of the play interview. *Ment. Hyg.*, 1939, **23**, 49-69.
8. FITE, M. D. Aggressive behavior in young children and children's attitudes toward aggression. *Genet. Psychol. Monog.*, 1940, **22**, 151-319.
9. GREEN, E. H. Group play and quarreling among pre-school children. *Child Devel.*, 1933, **4**, 302-307.
10. LEVY, D. M. Studies in sibling rivalry. *Res. Monog. Amer. Orthopsychiat. Assoc.*, 1937, No. 2.
11. JERSILD, A. T., & MARKEY, F. V. Conflicts between preschool children. *Child Devel. Monog.*, 1935, No. 21.
12. SEARS, R. R. Influence of methodological factors on doll-play performance. *Child Devel.*, 1947, **18**, 190-197.
13. WEISS-FRANKL, A. B. Play interviews with nursery school children. *Amer. J. Orthopsychiat.*, 1941, **11**, 33-40.

Department of Psychology
University of Louisville
Louisville 8, Kentucky

A STATISTICAL ANALYSIS OF AN ALUMNI SURVEY*

Testing and Instructional Research, Dearborn Public Schools

JOSEPH JACKSON

A. THE PROBLEM

Analyses of alumni surveys have in the past been largely, if not entirely, graphic or pictorial. This can be easily understood when one peruses the content of such surveys since in the main they are local, individual, and in parts totally subjective. Therefore to approach a diagnosis statistically, such data introduce a number of interesting problems. The paramount one is how much significance may be attached to the responses of alumni to particular questions and when should such responses have a bearing upon curriculum modification. If it were feasible to use some such approach, the problem of utilizing alumni surveys as factors in curriculum change would be greatly enhanced. As already inferred, the nature and the diversity of course content as well as the human differential seem to place a limitation upon the use of statistical techniques.

An intensive effort was made to evaluate by questionnaire the attitudes and experiences of Fordson High School graduates of the classes of 1938, 1941, 1944, 1947, and 1948. The purport of the questionnaire can be understood from the analyses compiled in Tables 1 and 2. A special emphasis in the survey was placed upon the curricular contributions of education to post-school experience which information also, served as a means of appraising present classroom practices. Out of 3213 questionnaires circulated, 808 returns were had, 67 per cent of which represented the most recent graduates or those of the classes of 1947 and of 1948. Incorrect addresses accounted largely for the mortality of returns among the preceding graduates.

B. STATISTICAL TREATMENT

Since the initial task in the statistical treatment of a survey is the reduction of data, formulas applicable through a common denominator are necessary. Two such formulas are available both being applicable to a percentage base. The first, and that used in this study, is known as the Standard Error of a Percentage in which case a critical ratio of 2.5 or more provides a significant relationship and the second is known as the Test of Significance and the

*Received in the Editorial Office on May 14, 1951.

Test Ratio (T). In the latter case fixed ratios are also utilized for indicating significance about the incidence of a given attribute for the first through the fifth per cent level of confidence. The Test of Significance is the more desired statistical formula for this study, however, it can only be applied to situations where parameter equivalents can be established. Since difficulty was experienced in establishing parameter bases, the standard error formula was utilized for the analyses of critical differences.

C. COMPILATION OF DATA

The information obtained was classified under 23 groupings of data. Table 1 portrays the number and per cent of responses common to each variable compared. The statistical calculations involved in each evaluation are presented in Table 2 which contains the differences between appropriate percentages for these variables. The standard errors of the differences, SE_d , the ratios of the differences to the standard errors of the differences, D/SE_d , and the chances of true differences greater than zero are also portrayed in Table 2. The value, 2.5 or higher, in the column D/SE_d eliminates the influence of chance factors upon the results.

Tables 3-6 provide by gender a rank order analyses of subject matter as it has been found by alumni to contribute to work pursued or to life adjustment in general. This tabular data therefore singles out various academic areas for departmental study and perhaps corrects many misgivings about the relative importance of the outcomes and goals embodied in such course content.

To enrich the analyses various data were combined wherever possible especially if the variables were related to some curricular implication. Such procedures permitted a broader interpretation of the statistical results and added a configurational approach to the study. In addition better analyses of departmental curricula as college preparatory, CP, general course, G, commercial, C, college-commercial, CC, industrial, I, and college-industrial, CI, were realized which was the primary intent of the study.

D. RESULTS OF COMPARISON OF VARIABLES BASED UPON 23 CLASSIFICATIONS OF DATA AS DISCLOSED BY FORDSON HIGH SCHOOL ALUMNI, OCTOBER, 1948

1. *Weighting*

The critical ratio of 9.3 indicates that the overall results of this survey are significantly weighted by responses of students with college preparatory and commercial course backgrounds. In addition the above weightings will be

TABLE 1
NUMBER AND PER CENT OF RESPONSES TO ITEMS OF ALUMNI QUESTIONNAIRE
FORDSON HIGH SCHOOL*

Variable (1)	No.	%	vs.	Variable (2)	No.	%
1. Weighting:						
College Preparatory	264	32.7		Commercial	231	28.6
Commercial	231	28.6		General	119	14.7
General	119	14.7		Industrial	98	12.1
College Prep. & Commercial	495	61.3		Other	313	38.7
Class '48	334	41.4		Class '47	210	26.0
Class '47-'48	554	67.4		Other	254	32.6
Female	469	58.0		Male	339	42.0
2. Subsequent Education:						
(CP) College Attendance	137	51.9		Non-attendance	127	48.1
(G) None	89	74.8		Yes	30	25.2
(C) None	196	84.9		Yes	35	15.1
(CC) College Attendance	14	41.2		Non-attendance	20	58.8
(I) None	60	61.2		Yes	38	38.8
(CI) Yes	37	66.1		Non-attendance	19	33.9
(CI) College Attendance	32	86.5		Technical School	5	13.5
3. Subjects Greatest Value to Work:						
Boys—English	157	19.7		Mathematics	135	17.0
English	157	19.7		Science	121	15.2
Girls—Commercial Subjects	520	44.0		Eng., Math., Social		
Typing	241	20.4		Study, Social Math.	431	36.5
				Shorthand	124	10.5
4. Subjects Least Value to Work:						
Boys—Social Studies	104	24.4		Science	51	12.0
Girls—Science	92	16.0		Geometry	44	7.67
5. Subjects Greatest Value to Life:						
Boys—English	174	27.2		Social Studies	126	19.7
Girls—English	204	22.8		Social Studies	165	18.6
6. Subjects Least Value to Life:						
Boys—Social Studies	47	20.7		Chemistry & Biology	22	9.69
Girls—Mathematics	57	19.45		History	19	6.48
7. Recreation, First and Second Choice:						
Boys—All sports	515	83.3		Dance, Music, Read.,		
Swimming	110	17.8		& Writing	103	16.7
				Baseball	78	12.6
8. Recreations, First and Second Choice:						
Girls—All Sports	691	75.6		Dance, Music, Read.,		
				& Writing	223	24.4
Swimming	267	29.2		Dancing	104	11.4
9. Suggestions for Making School More Helpful for:						
Undergraduates—						
Motivate Interest,				Guidance &		
Stress Other Subjects,				Counseling	224	25.8
Demand More Intensive						
Work	413	47.6		Other	399	46.0
Guidance, Counsel &						
Motivate Interest	469	54.0				

*CP—College Preparatory Major
G —General Course Major
C —Commercial Major

CC—College—Commercial Major
I —Industrial Major
CI —College—Industrial Major

TABLE 1 (continued)

Variable (1)	No.	%	vs.	Variable (2)	No.	%
10. For Alumni and Drop Outs: Guidance & Counsel, Job Placement	183	41.5		Alumni Meetings	82	18.6
11. School Could Have Helped Me More: Counsel, Guidance, Occupa- tional Information and Obtain Job	807	51.4		Preparation for Life and Solve Person- ality Problems	388	24.7
Preparation for Life and Solve Personality Problems	388	24.7		Direct Job Training	328	20.9
12. Services for Graduates and Drop Outs: Night School, Spare Time Study, School A Center for Educational Activities	1081	59.4		Guidance & Counsel, Find Jobs	722	39.8
13. Means of Obtaining Present Job: Self	238	48.5		School	101	20.5
School	101	20.5		Friend	57	11.6
14. Means of Obtaining First Job: Self	238	44.7		School	154	28.9
School	154	28.9		Friend	52	9.8
School	154	28.9		Home	29	5.4
Friend	52	9.8		Home	29	5.4
15. Location of Employment—Present Job: City	209	42.4		Detroit	222	45.0
City	209	42.4		Other	62	12.6
16. Location of Employment—First Job: City	238	44.0		Detroit	262	48.4
City	238	44.0		Other	41	7.6
17. Employment of Graduates—Present Job: (CP) Commercial	61	40.1		Industrial	24	15.8
Commercial	61	40.1		Service Occupations	44	28.9
Industrial	24	15.8		Service Occupations	44	28.9
(G) Commercial	32	56.1		Industrial	12	21.1
(C) Commercial	131	89.1		Service Occupations	10	6.8
(CC) Commercial	17	73.9		Other	6	26.1
(I) Industrial	51	69.9		Commercial	10	13.7
(CI) Industrial	19	54.3		Service Occupations	8	22.9
18. Employment of Graduates—First Job: (CP) Commercial	69	40.6		Service Occupations	59	34.7
Industrial	21	12.4		Service Occupations	59	34.7
Professional	21	12.4		Service Occupations	59	34.7
(G) Commercial	39	39.0		Service Occupations	26	26.0
Service Occupations	26	26.0		Industrial	12	12.0
(C) Commercial	147	80.8		Service Occupations	27	14.8
(I) Industrial	42	63.6		Service Occupations	13	19.7
Industrial	42	63.6		Other	24	36.4
(CI) Industrial	21	60.0		Other	14	40.0
19. Reasons for Unemployment: Attend School, Home Duties, And Other Reasons	279	79.7		Lack of Training and Experience, No Avail- able Jobs	71	20.3
Other Reasons	300	85.7		No Jobs, Lack of Training	50	14.3

TABLE 1 (continued)

Variable (1)	No.	%	vs.	Variable (2)	No.	%
20. Frequency and Dominant Occupational Choice at Graduation:						
(CP) Yes (College Major)	166	67.5	No		80	32.5
(G) Yes (Commercial)	57	54.3	No		48	45.7
(C) Yes (Commercial)	171	79.2	No		45	20.8
(CC) Yes (Comm. & College Major)	27	84.4	No		5	15.6
(I) Yes (Industrial)	49	68.1	No		23	31.9
(CI) Yes (College Major)	39	72.2	No		15	27.8
21. Extent of Job Satisfaction:						
Yes	512	81.9	No		113	18.1
Interest and Occupation Related	359	57.4	Other		266	42.6
Changed Original Occupational Plan	62	9.9	Interest & Occupation Related		359	57.4
22. Most Help in Making Educational Plans:						
Parent	333	26.9	Course		303	24.5
Teacher	139	11.2	Counselor		158	12.8
Parent	333	26.9	Counselor		158	12.8
Course, Counselor, Teacher	600	48.5	Parent		333	26.9
23. Classroom Work May Be More Effective:						
(CP) Motivate Interest	143	34.0	Modify Curriculum		70	16.6
Motivation, Democratic Classroom Procedures, Intense Work	233	55.3	Modify Curriculum, Add Courses, Job Training		115	27.3
(G) Motivation, Democratic Classroom Procedures, Intense Work	102	67.5	Other		49	32.5
(C) Motivation, Democratic Classroom Procedures, Intense Work	169	54.5	Modify Curriculum, Add Courses, Job Training		74	23.9
(I-CI) Motivation, Democratic Classroom Procedures, Intense Work	106	52.7	Modify Curriculum, Add Courses, Job Training		65	32.0

especially influenced by graduates of the Class of 1948. The critical ratio in the latter case is 6.67.

Since a part of the weighting is a result of commercial course majors, it thereby introduces a sex factor into the forthcoming analyses. The critical ratio between girl-boy weightings is 6.5 or considerably above the expected minimum of 2.5. Female graduates because of their preponderance among commercial students will color the total results in this study as well as any results concerned with the commercial department.

2. Subsequent Education

Of the 264 college preparatory students responding, 51.9 per cent continued college. From the standpoint of students specializing, an unusually high

TABLE 2

DIFFERENCES BETWEEN VARIABLES BASED UPON TWENTY-THREE CLASSIFICATIONS OF DATA
FORDSON HIGH SCHOOL ALUMNI SURVEY

(1)	Variables compared (2)	Differences between percentages 1 vs. 2	SE _d	D		Chances in 100 of a true difference greater than zero
				SE _d	D	
1.	Weighting:					
	College Prep. vs. Commercial	4.1	7.25	.57		64
	Commercial vs. General	13.9	6.39	2.18		93
	General vs. Industrial	2.6	5.36	.49		63
	College Prep. & Commercial vs. Other	22.6	2.43	9.3		100
	Class '48 vs. Class '47	15.4	2.31	6.67		100
	Class '47-'48 vs. Other	34.8	1.65	21.09		100
	Female vs. Male	16.0	2.46	6.5		100
2.	Subsequent Education:					
	(CP) College Attendance vs. Non-attendance	3.8	4.35	.87		72
	(G) None vs. Yes	49.6	3.98	12.46		100
	(C) None vs. Yes	69.8	2.36	29.58		100
	(CC) College Attendance vs. Non-attendance		17.6	8.44	2.09	92
	(I) None vs. Yes	22.4	3.46	6.47		100
	(CI) Yes vs. Non-attendance	32.2	6.33	5.09		100
	(CI) College Attendance vs. Technical School	73.0	5.62	12.99		100
3.	Subjects Greatest Value to Work:					
	Boys—English vs. Mathematics	2.7	1.94	1.39		83
	English vs. Science	4.5	1.9	2.37		95
	Girls—Commercial Subjects vs. English, Math., Social Study, Social Mathematics	7.5	2.01	3.73		99
	Typing vs. Shorthand	9.9	1.47	6.73		100
4.	Subjects Least Value to Work:					
	Boys—Social Studies vs. Science	12.4	2.61	4.75		100
	Girls—Science vs. Geometry	8.36	1.89	4.42		100
5.	Subjects Greatest Value to Life:					
	Boys—English vs. Social Studies	7.5	2.36	3.18		98
	Girls—English vs. Social Studies	4.2	1.92	2.19		93
6.	Subjects Least Value to Life:					
	Boys—Social Studies vs. Chemistry & Biology	11.0	3.33	3.3		99
	Girls—Mathematics vs. History	12.97	2.73	4.75		100
7.	Recreation, First and Second Choice:					
	Boys—Sports vs. Dance, Music, Reading and Writing	66.6	2.12	31.4		100
	Swimming vs. Baseball	5.2	2.07	2.5		95
8.	Girls—Sports vs. Dance, Music, Reading, Writing	51.2	2.46	20.8		100
	Swimming vs. Dancing	17.8	1.84	9.67		100
9.	Suggestions for Making School More Helpful for:					
	Undergraduates—					
	Motivate Interest, Stress Other Subjects, Demand More Intense Work vs. Guidance & Counseling	21.8	2.25	9.69		100
	Guidance & Counseling, Motivate Interest vs. Other	8.0	1.97	4.06		100

TABLE 2 (continued)

(1)	Variables compared (2)	Differences between percentages 1 vs. 2	SE _d	D		Chances in 100 of a true differ- ence greater than zero
				SE _d	SE _d	
10.	Alumni and Drop Outs: Guidance & Counseling, Job Placements vs. Alumni Meetings	22.9	2.99	7.66	100	
11.	School Could Have Helped Me More: Counseling & Guidance, Occupational Informa- tion and Obtain Job vs. Solve Personality Problems and Preparation for Life	26.7	1.67	15.99	100	
	Solve Personality Problems and Preparation For Life vs. Job Training	3.8	1.5	2.53	95	
12.	Services for Graduates and Drop Outs: Night School, Spare Time Study & School A Center for Educational Activities vs. Guidance and Counseling, Find Jobs	19.6	1.62	12.1	100	
13.	Means of Obtaining Present Job: Self vs. School	28.0	2.9	9.66	100	
	School vs. Friend	8.9	2.33	3.82	99	
14.	Means of Obtaining First Job: Self vs. School	15.8	2.91	5.43	100	
	School vs. Friend	19.1	2.35	8.13	100	
	School vs. Home	23.5	2.19	10.73	100	
	Friend vs. Home	4.4	2.54	1.73	88	
15.	Location of Employment—Present Job: City vs. Detroit	2.6	3.16	.82	71	
	City vs. Other	29.8	2.68	11.1	100	
16.	Location of Employment—First Job: City vs. Detroit	4.4	3.03	1.45	84	
	City vs. Other	36.4	2.42	15.0	100	
17.	Employment of Graduates—Present Job: (CP) Commercial vs. Industrial	24.3	4.95	4.91	100	
	Commercial vs. Service Occupations	11.2	5.42	2.07	92	
	Industrial vs. Service Occupations		13.1	4.72	2.78	97
	(G) Commercial vs. Industrial	35.0	8.51	4.11	100	
	(C) Commercial vs. Service Occupations	82.3	3.34	24.6	100	
	(CC) Commercial vs. Other	47.8	9.1	5.25	100	
	(I) Industrial vs. Commercial	56.2	6.71	8.38	100	
	(CI) Industrial vs. Service Occupations	31.4	11.02	2.85	97	
18.	Employment of Graduates—First Job: (CP) Commercial vs. Service Occupations	5.9	5.25	1.13	77	
	Industrial vs. Service Occupations		22.3	4.44	5.02	100
	Professional vs. Service Occupations		22.3	4.44	5.02	100
	(G) Commercial vs. Service Occupations	13.0	6.56	1.98	91	
	Service Occupations vs. Industrial	14.0	5.48	2.55	95	
	(C) Commercial vs. Service Occupations	66.0	4.4	15.0	100	
	(I) Industrial vs. Service Occupations	43.9	7.7	5.7	100	
	Industrial vs. Other	27.2	5.92	4.59	100	
	(CI) Industrial vs. Other	20.0	8.27	2.42	95	

TABLE 2 (continued)

Variables compared		Differences between percentages	SE _d	D	Chances in 100 of a true difference greater than zero	
(1)	(2)	1 vs. 2		SE _d		
19.	Reasons for Unemployment: Attend School, Home Duties & Other Reasons vs. Lack of Training, Lack of Experience, No Available Jobs	59.4	2.15	27.63	100	
	Other Reasons vs. No Jobs, Lack of Training	71.4	2.65	26.9	100	
20.	Occupational Choice at Graduation:					
	(CP) Yes vs. No	35.0	2.99	11.71	100	
	(G) Yes vs. No	8.6	4.88	1.76	88	
	(C) Yes vs. No	58.4	2.76	21.15	100	
	(CC) Yes vs. No	68.8	6.42	10.72	100	
	(I) Yes vs. No	36.2	5.5	6.58	100	
	(CI) Yes vs. No	44.4	6.2	7.16	100	
21.	Extent of Job Satisfaction:					
	Yes vs. No	63.8	1.43	44.6	100	
	Interest & Occupation Related vs. Other Changed Original Occupational Plan vs. Interest & Occupation Related	14.8	1.9	7.79	100	
22.	Most Help in Making Educational Plans:	47.5	2.39	19.87	100	
	Parent vs. Course	2.4	1.76	1.36	82	
	Teacher vs. Counselor		1.6	1.3	1.23	80
	Parent vs. Counselor	14.1	1.6	8.81	100	
	Course, Counselor, Teacher vs. Parent	21.6	1.9	11.37	100	
23.	Classroom Work May Be More Effective:					
	(CP) Motivate Interest vs. Modify Curriculum	17.4	2.94	5.92	100	
	Motivation, Democratic Classroom Procedures, Intense Work vs. Modify Curriculum, Additional Courses, Job Training	28.0	3.26	8.59	100	
	(G) Motivation, Democratic Classroom Procedures, Intense Work vs. Other	35.0	3.8	9.21	100	
	(C) Motivation, Democratic Classroom Procedures, Intense Work vs. Modify Curriculum, Add Courses, Job Training	30.6	3.72	8.23	100	
	(I-CI) Motivation, Democratic Classroom Procedures, Intense Work vs. Modify Curriculum, Add Courses, Job Training	20.7	4.79	4.32	100	

mortality exists among alumni who do not continue higher education. The critical ratio of .87 bears this out strongly. What then are the post-school problems of the 48.1 per cent who do not go on to college? This mortality suggests perhaps that a follow-up study is needed as well as a greater guidance and counseling consciousness during the high school tenure of college preparatory students.

Neither general course nor commercial course majors resort to additional schooling upon graduation. The high critical ratios of 12.46 and 29.58 respectively indicate practically unanimous agreement in this matter. In the case of commercial course graduates, it appears that the high percentage of related job placement designates an adequate vocational preparation in high school.

Do college-commercial majors tend to attend college in large numbers? Of the thirty-four responses received, only fourteen went on to college, an insignificant number as shown by the critical ratio of 2.09.

TABLE 3
SUMMARY OF SUBJECTS BY RANK ORDER HAVING GREATEST VALUE TO WORK PURSUED
FORDSON HIGH SCHOOL—OCTOBER 1948

Boys	Rank		Girls	Rank	
English	157	1	Typing	241	1
Mathematics	135	2	English	238	2
Science	121	3	Shorthand	124	3
Shop and Drawing	73	4	Social Studies	88	4
Social Studies	54	5	Mathematics	83	5
Typing	17	6	Bookkeeping	51	6
Physical Education	16	7.5	Office Practice	49	7
Public Speaking	16	7.5	Science	42	8
Music	12	9	Social Math	22	9
Other	195	10	Filing	21	10
			Comptometer	19	11
Responses	796		Business English	15	12
			Other	188	13
			Responses	1181	

Of the 98 industrial course graduates responding, 60 or 61.2 per cent terminated their education in high school. The significance ratio of 6.47 bears out the lack of educational follow-up. Of the 38.8 per cent reporting subsequent education about half attend college and half attend a technical school.

College-industrial majors flock in significantly high number to advanced education. This trend is illustrated by the critical ratio of 5.09. Of the 37 students seeking advanced education, 32 or 86.5 per cent attend college. It should be noted that the incidence of college attendance is better than it is upon a college preparatory course.

3. *Subjects Greatest Value to Work*

Boys rank English, mathematics, and science in that order as being of greatest value to work. There are no significant differences between these

rank orders. The differences however between any of the three subject fields and the remaining subjects ranked for preference in Table 3 are significant.

Girls rank typing, English, and shorthand respectively as being of greatest value to work pursued. No significant difference exists between typing and English, however, the critical ratio of 6.73 designates a strong preference for typing when compared to shorthand. A similar critical ratio favors English over shorthand. The difference between any of the first three subjects and any other given in Table 3 would probably be significant.

In the case of girls, it should be stated that the value of commercial subjects significantly outweighs the value of the academic when compared in the overall picture of Table 3. The ratio of difference is 3.73 or considerably above the minimum of 2.5.

The above summaries are compatible with the weightings. Since the weightings are of a college preparatory and of a commercial course nature, they would likewise be consistent with the academic and vocational needs of each group. Since females are dominantly the majority in this survey, their commercial backgrounds are reflected here.

English is the dominant choice of boys and girls as far as value of subject matter and its relationship to work are concerned.

4. *Subjects Least Value to Work*

Boys feel that social studies are of least value to work pursued with the field of science being second in this respect. The rank differences shown in Table 4 between these choices are significant as shown by the critical ratio of 4.75.

The girls on the other hand state that science and social studies are of about equal ineffectiveness when analyzed through the functional requirements

TABLE 4
SUMMARY OF SUBJECTS BY RANK ORDER HAVING LEAST VALUE TO WORK PURSUED
FORDSON HIGH SCHOOL—OCTOBER 1948

Boys		Girls	
	Rank		Rank
Social Studies	104 1	Science	92 1
Science	51 2	Social Studies	87 2
English & Literature	43 3	Geometry	44 3
Geometry & Mathematics	37 4	Physical Education	34 4
Latin	15 5	Bookkeeping	28 5
Shop and Drawing	12 6	Filing	24 6
Other	164 7	Shorthand	20 7
		Latin	16 8
Responses	426	Other	229 9
		Responses	574

of various course backgrounds. The remaining differences are significantly apart as shown by the science-geometry ratio of 4.42.

It is worthy to note that boys dominantly college preparatory in schooling and girls dominantly commercial in vocational pursuits agree on the lack of subject matter relationship to work pursued after high school graduation.

5. *Subjects Greatest Value to Life*

English and social studies rank first and second as choices of subjects having the greatest value to life in general of both boys and girls. See Table 5 for the summary. English outweighs social studies significantly in importance

TABLE 5
SUMMARY OF SUBJECTS BY RANK ORDER HAVING GREATEST VALUE TO LIFE IN GENERAL
FORDSON HIGH SCHOOL—OCTOBER 1948

Boys	Rank	Girls	Rank
English	174 1	English	204 1
Social Studies	126 2	Social Studies	165 2
Mathematics	75 3	Homemaking, Sewing	93 3
Physical Education	50 4	Physical Education	61 4
Science	48 5	Science	49 5
Music	15 6.5	Mathematics	45 6
Public Speaking	15 6.5	Music	24 7
Shop and Drawing	9 8	Typing	23 8
Clubs	7 9	Clubs	19 9
Other	121 10	Art	14 10
Responses	640	Other	197 11
		Responses	894

as a subject field of great value to boys in post-school life. The significance ratio is 3.18. Girls, however, do not distinguish significantly between English and social studies since in a similar comparison the critical ratio of 2.19 is removed from the minimum criterion ratio of 2.5. The differences on behalf of the choices of boys as well as girls between English or social studies and other subjects of value to life are significant.

6. *Subjects Least Value to Life*

As far as the life needs of boys are concerned social studies seem to be of least value. There were 47 replies out of 227 in this direction. Refer to Table 6 for rank of difference. Girls show a similar trend of thinking in relation to mathematics. Chemistry and biology are second choices for boys and history a second choice for girls. While the difference between the first and second choice of boys and the first and second choice of girls are significant, the results should be treated from the number of survey returns which make

TABLE 6
SUMMARY OF SUBJECTS BY RANK ORDER HAVING LEAST VALUE TO LIFE IN GENERAL
FORDSON HIGH SCHOOL—OCTOBER 1948

Boys		Rank		Girls		Rank	
Social Studies	47	1		Mathematics	57	1	
Chemistry and Biology	22	2		History	19	2	
Geometry	15	3		Biology	14	3	
Literature	13	4		Typing	13	4	
Shop and Drawing	12	5.5		Shorthand	11	5.5	
Latin	12	5.5		Bookkeeping	11	5.5	
Physical Education	9	7		Physical Education	10	7.5	
Art	7	8		Horticulture	10	7.5	
Other	90	9		Other	148	9	
Responses	227			Responses	293		

for the above analysis. Furthermore, the criticisms of subjects of least value should likewise be compared to those having greatest value to life so that a configurational approach may be made.

In the evaluation of subject matter as offered in high school and its place in the after school life of an alumnus, considerable study should be given the curricular aspects of this phase of the survey. Boys as well as girls acting independently have provided information which from their point of view increases the functional benefits of certain subject-matter fields. In such instances it is important that a false sense of security does not accrue since the returns mean keeping abreast of current needs. In other areas where subject matter fields do not seem to fit the needs of students, perhaps curricular development will be in order.

7. Recreation: Boys

The first and second choices of recreational interests of boys are overwhelmingly of a sporting nature. When such choices are compared to dancing, music, reading, and writing, the critical ratio is 31.4 results. Note the departure of this ratio from the minimum significance test of 2.5.

Of the individual choices preferred, what is the favorite sporting activity of boys? When the two top choices are compared, the ratio of 2.5 indicates that swimming is preferred over baseball although both are popular. This is a healthy trend, since the former is an active sport while the latter is definitely of the arm-chair and seasonal variety to a goodly portion of our alumni. For future trends, what are the local recreational needs? Is baseball the so-called "national" sport or is its popularity regressing?

8. *Recreation: Girls*

Girls likewise favor sports over dancing and reading with a similar degree of bias as do the boys. The critical ratio of 20.8 portrays the degree of preference for sports as compared to dancing, music, or reading.

The individual activities preferred by girls are swimming first, dancing second. The critical ratio of 9.67 designates the degree to which swimming is chosen.

9. *Suggestions for Making School More Helpful for: Undergraduates*

Alumni suggest that increased motivational approaches to classroom problems would make school more helpful for undergraduates. The survey points out that motivation through interest and the stressing of assignments are significantly more important than is guidance or counseling. The critical ratio is 9.69 in such a comparison.

The need for guidance and counseling when compared to other undergraduates' needs likewise shows a high critical ratio of 4.06 in favor of the former variable.

10. *Alumni and Drop Outs*

As far as their own needs are evaluated, alumni feel that guidance, counseling, and job placements are much more important than organized alumni meetings. The critical ratio of 7.66 indicates this feeling to a significant degree.

11. *School Could Have Helped Me More*

The dominant need of alumni appears to be of a guidance and counseling nature. If *guidance and counseling, the need for occupational information and obtaining a job* are combined as one variable and are then compared to the combination variable of *solving personality problems* and the *need of preparation for life*, the critical ratio is 15.99.

The solving of *personality problems* when combined with the need for *preparation for life* is compared to the variable of *job training* yields a critical ratio of 2.53. Following guidance and counseling, personality and life adjustment posit the next important need of alumni.

12. *Services for Graduates and Drop Outs*

The services for graduates appear to be of an educational nature. Night school, spare time study, and the use of the school as a center for educational activities seem to be more necessary than guidance, counseling, or the finding

of jobs. The critical ratio of 12.1 emphasizes preponderantly the educational aspects over the guidance ones.

This questionnaire did not attempt to evaluate the specific part-time educational needs of alumni. It merely inquired of services that could be provided for graduates and drop outs. Perhaps many of those replying are merely suggesting educational outlets rather than planning evening school work. To discern between such questions, a follow-up circular would have to be used to determine the nature and extent of part-time or evening school interests.

13. *Means of Obtaining: Present Job*

The self is the principal means of obtaining the present job when compared to the influence of the school or of a friend in such matters. The self-school influence as shown by the critical ratio of 9.66 designates the degree to which the self is active.

The influence of the school is likewise significantly heavier than is the aid of a friend in procuring the present job for alumni. The critical ratio in the latter instance is 3.82.

14. *First Job*

The first job is most generally obtained by the self according to Fordson High School alumni. The two most common methods, namely self and school, provide a critical ratio of 5.43 which is significantly on the side of own job procurement.

The school is much more important than is a friend in securing a job. The influence of the home is practically negligible when compared to the influence of the self, the school, or of a friend in obtaining the first job for an alumnus. It seems then that the home should be brought into the job placement and perhaps into the entire guidance picture more concretely.

15. *Location of Employment*

Graduates are employed at their present or at their first job to an equal degree in the City of Dearborn or in Detroit. Other communities provide an insignificant amount of employment for our graduates. If the latter is true then what may be done to provide additional vocational outlets for alumni especially in such industrial areas as the down-river district?

16. *Employment of Graduates*

In the light of course background, what vocational endeavors do Fordson High School graduates pursue? Do the replies tend to reflect the occupational

fate of 48 per cent of college preparatory students who do not matriculate at an institution of higher learning? Perhaps the following summaries may shed light on both of the above queries.

17. *Present Job*

College preparatory alumni employed at their present job tend to favor employment of a commercial type. Their second favorite type of employment is a service occupation such as laundry or dry cleaning work, etc. Statistically no significant differences exist as to either choice which is borne out by the critical ratio of 2.07.

College preparatory students significantly favor commercial type of employment when compared to an industrial type. The female weighting should be recognized. The critical difference in this case is 4.91. College preparatory students likewise favor service occupations over industrial ones. The critical ratio in the latter case is 2.78.

General course students predominantly follow a commercial occupation. The commercial-industrial type of employment difference of 4.11 is definitely to the advantage of the former. Here again the female sex factor weights the result.

Alumni with commercial majors dominate their own field of employment. When commercial-service employment is studied as a present job the critical ratio of 24.6 forecasts the degree to which students with commercial backgrounds stay within their own area of work.

College-commercial majors prefer employment commercially than otherwise. The critical difference in such a comparison is 5.25. Only 23 responses to this question were received.

Industrial course graduates tend to stay within their own area of specialization. When compared to commercial types of employment, the critical ratio of 8.38 demonstrates a heavy preference for industrial work.

College-industrial people favor industrial work to service occupations. The critical difference of 2.85 weights it significantly in that direction. Only 35 responses were received to this question.

18. *First Job*

College preparatory students as in the case of the present job tend toward a commercial type or a service type occupation. The critical ratio of 1.13 points to about an equal dispersion of employment within these two fields. Industrial employment of college preparatory majors is of secondary importance as far as employment is concerned.

It seems that perhaps by both nature and nurture, a college preparatory student is predisposed for white collar work. Since this appears to be the case, maximum preparation of this type should be in order for the 48 per cent of college preparatory students terminating their specialization at the secondary school level. One hundred seventy college preparatory graduates replied to this question.

General course students accept commercial or service type of employment in preference to industrial for their initial entry into employment. One hundred general course majors responded to this question. On the basis of first or of present employment, a general course should incorporate sufficient commercial training to provide entry into commercial fields of employment. It appears from all preceding data that from a placement angle and from a guidance angle the so-called service occupations should be surveyed and adapted to the educational needs of general course students.

Commercial course graduates without further experience are placed directly into their own field of specialization. Service occupations for initial employment are a poor second choice as evidenced by the critical ratio of 15.0.

Industrial course majors start to work within their field of specialization to a marked degree.

Industrial-service occupations have a critical difference of 5.7 in the direction of the former variable while industrial-other type of employment show a critical difference of 4.59 also in favor of industrial work. College-industrial graduates do not show any fixed employment pattern. Industrial work is favored, however, the returns of the survey are too few to affect a thorough diagnosis.

19. *Reasons for Unemployment*

Assuming that the lack of training, lack of experience, and scarcity of jobs are major factors in unemployment which are then compared to such reasons as attending school, home duties, etc., that make for employment, what then is our scene? In such a comparison the unusually high critical ratios of 27.63 and 26.9 as given in Table 2 dispel any concept of unemployment.

20. *Frequency and Dominant Occupational Choice at Graduation*

The summary about this area of the survey should be studied in relation to foregoing data. Such critical items as subsequent education, employment, guidance and counseling are inherent in the deeper aspects of this particular problem. The following factor should be retained. Frequency and dominant occupational choice at time of high school graduation does not necessarily

mean consistency in post-school follow-up. Data already presented on subsequent education bear this out.

College preparatory students plan college major pursuits. However, the data of Table 2 indicate fallacious planning in 48.1 per cent of such cases. When one considers also that of the 51.9 per cent of college preparatory graduates matriculating in an institution of higher learning, a goodly portion will not receive a bachelor's degree, one sees the complexity of problems inherent in this area of vocational guidance.

General course students flounder noticeably from the standpoint of vocational goals at time of graduation. The critical ratio of 1.76 is insignificant as far as definite occupational planning concerns this group. It is of utmost importance to note that general course majors as a group are without a vocational objective at time of graduation when compared to graduates of other courses. From this analysis it appears that the general curriculum should probably either be reorganized, dissolved, or incorporated into another major, preferably commercial in nature since about half of general course students drift into such occupations.

Commercial students adhere to this area of specialization and placement. Previous data have reiterated this inference through placements and lack of subsequent schooling for these specialists. It appears that the high school commercial course attains maximum functional capacity within the occupational life of such majors. From this respect it is unique.

College commercial majors are likewise well defined in their occupational objectives at time of graduation as shown by the critical ratio of 10.72. The findings of this paragraph dovetail with the analysis previously made of subsequent education and the subjects of greatest value to work for this group.

College-industrial students show a similar trend with college follow-up being the significant trend. Industrial course majors plan occupational outlets within their area of specialization. The critical ratio of 6.58 indicates then that vocational pursuits out of this field of work are rare and insignificant. Seventy-four responses were analyzed for the latter summary.

21. *Extent of Job Satisfaction*

Fordson High School alumni seem to be unusually well adjusted in their present occupational pursuits. The critical ratio in this case is 44.6 which is considerably above the minimum 2.5 ratio of significance. What is the principal reason for this degree of occupational adjustment? In answer, the alumni state that their interests are closely allied with their vocational outlets.

To what extent does change occur which offsets the original vocational

objective of alumni? Since the interests of alumni are closely related to their occupational endeavors, change of occupation occurs in an insignificant degree. The critical ratio of 19.87 in such a comparison reverts strongly toward the relationship of interests and occupations.

22. *Most Help in Making Educational Plans*

The low critical ratio of 1.36 indicates that parental guidance and school course-content materials competed about evenly in formulating the educational plans of former students. These two influences are unique in this respect while the influence of the teacher or counselor are somewhat less effective. The parent for example is significantly mentioned in the formulation of educational plans when compared to the counselor. The critical ratio of 8.81 designates marked effectiveness of parental interests and attention in matters of educational planning of their children. A similar critical ratio exists on behalf of course content over the influence in such matters of friends, counselors, or employers.

How does the school team of course content, counselor, and teacher function in the educational planning of graduates when compared to the influence of the parent? The critical ratio of 11.37 definitely skews the educational planning of alumni in the direction of the school.

23. *Classroom Work May Be More Effective*

a. *College preparatory students.* Four hundred twenty-one responses were received from this group of former students. How critical are the differences between the two major suggestions for making classroom work more effective? In such a comparison the desire for more motivation within the classroom is significantly much more important than is direct curriculum modification. The critical ratio in favor of the former variable is 5.92 or considerably above the minimum critical difference of 2.5. If on the other hand such commonly related criticisms as motivation of interest, the need for democratic classroom procedures, and the requirement of intense work are grouped as one variable and compared to the grouped variable of modify curriculum, provide additional courses and provide job training, the critical ratio of 8.59 suggests a strong demand for improved motivational approaches to academic situations.

b. *General course students.* This group of students tells a similar story to that of the college preparatory group except its wants in the direction of motivational needs are more intense. The critical ratio of 9.21 illustrates the need for change in the direction of motivation of interest, more democratic classroom procedures and more intense study when compared to sundry other criticisms.

c. *Commercial course students.* Commercial course graduates seem to be rather well satisfied with the curriculum and job training. They express through significant numbers a request for more variation in classroom procedures as do college preparatory and general course graduates. The critical ratio for their preference is 8.23 which is considerably above the minimum test of significance of 2.5.

d. *Industrial and college industrial students.* These two groups of students likewise express a need for improved classroom motivational techniques as did alumni who graduated with the backgrounds summarized above. Job training, new courses or curriculum modification are not as necessary as are improved operational procedures within the classroom. The critical ratio of 4.32 in such an analysis suggest that such an expression is conditioned by other than chance factors.

It appears from the above analysis that graduates of the major high school courses are well agreed upon the need for change in motivational techniques allied with classroom pedagogy. Because of the breadth of the general survey, 808 responses, an extended study on a departmental basis should be made of the issues involved. Change in the direction requested can hardly be affected by diversity of procedure on behalf of few instructors. It appears that the criticisms are more thorough and require departmental approaches. Such a study group should not concern itself alone with the findings of this latter summary. It should correlate other factors as least and most valued subject areas, sex preferences, relationship of course to post-school experience, and the like. Such an approach should produce results of benefit to the student. That is its purpose.

E. CONCLUSIONS

A statistical analysis of data such as compiled for this study offers an opportunity to appraise the importance attached to responses which might have a bearing upon such delicate processes as curriculum revision. Unusually desirable is the amalgamation of variables into areas of curriculum responsibility which not only make minute analyses possible, but also designate contributions to departmental influences as they bear upon educational outcomes. This has been found especially true in the fields of guidance, counseling, subjects of least and greatest value, and recommendations for more effective classroom work. As a by-product administrative implication becomes more readily accessible and scientific.

Greater application of statistics should be possible to surveys of the alumni type. For example, the Test of Significance through the Test Ratio (T) produced two problems. First, the inability to agree upon what constitutes a

suitable parameter base for such minority electives as homemaking, Latin, speech, modern living, and the like. Second, the small percentage of alumni returns about these minority electives appear to delimit the use of the Test Ratio (T) inasmuch as the parameter values exceed the sampling thereby providing negative tests of insignificance.

Department of Testing and Instructional Research
5757 Neckel Avenue
Dearborn, Michigan

PATTERNS OF VOCATIONAL INTEREST DEVELOPMENT AND ACTUAL JOB CHOICE*¹

Department of Psychology, Syracuse University

JOSEPH L. NORTON

A. THE NEED FOR THIS STUDY

The importance of the choice of a life occupation is self-evident. Clearly, the development leading to vocational choice should be well understood if youth are to be helped to intelligent choices. However, review of the studies in the field of interest growth reveals many gaps in our knowledge. There is a paucity of research attempting to trace the development of vocational preference of individuals from the time of their beginning to their successful satisfaction in work, or to their discard for another interest or job. Furthermore, much of what is available may be seriously criticized because of the inadequacy of the research methods employed. The first part of this article will be devoted to a survey of past studies of the development of vocational interests, with attention devoted first to the findings, and secondly to the methodology. In both areas the need for further study is evident. The remainder of the paper will report an original investigation.

1. *Results of Earlier Studies*

a. Developmental trends in interest growth. Although the literature on vocational interests has been previously surveyed by Strong (15, 1943) and others (1, 1944; 2, 1944; 4, 1931; and 16, 1947), there have been no critical reviews which have attempted to emphasize the research on the *developmental trends* in interest growth. This review will summarize available facts regarding this development and the factors involved. The major findings may be stated as follows:

1. Vocational choices are made (preferences expressed) at all ages studied from three up. However, at all ages studied, some subjects failed to express a preference. As many as 55 per cent of a group have been reported to give no choice.

*Received in the Editorial Office on May 16, 1951.

¹Based on a portion of a dissertation submitted to Syracuse University in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The author gratefully acknowledges the helpful guidance received from Dr. Raymond G. Kuhlén.

2. No overall increase of the number making a choice is found with increased age; the majority are willing at most ages to name a preference.

3. Most writers report there is no general "improvement" in the type of choice with age, in terms of fitting the national or local figures of actual job possibilities. There is some disagreement with this finding, which has been blamed on methodology.

4. Sex differences reveal that girls appear to make choices earlier, and are more homogeneous in their choices.

5. In general, interests tend to become slightly more sedentary with age.

6. There is little improvement, but some change with age, of reasons given by subjects as influencing their preferences.

7. Interests reported as held earliest tend to hold over better into the work pattern.

8. Changes in interest may perhaps be brought about by definite, extended effort, but there is only contradictory evidence as to how far counseling, guidance, or "orientation classes" can go in effecting such changes.

Apparently the literature contains no tabulations of job choices of individuals through the years, nor of how interests waxed and waned, nor of just why one interest was lost and another acquired. No attention has been paid to the possibility that several interests are held simultaneously. There is virtually no carry-over of the study of vocational interest growth into the adult years; what there is appears to be concentrated in the field of teaching. The dearth of developmental age data on the emergence of vocational interests and preferences indicates definite need for such research.

b. Factors influencing interest growth. Besides yielding the meager data just indicated about trends in interest growth, studies of interests have attacked the problem of what influenced that growth. These have revealed that:

1. There is no agreement among the reports as to the *most* important influences. For example, 10 different influences are reported as "most important" by 27 groups investigated in 19 studies.

2. The *relative* importance of influences has not been adequately studied.

3. Although some studies report the family as the *least* important, there is considerable evidence for the effect of their families on the subjects.

4. Inference from stated preferences shows there has been considerable impact of the social recognition (prestige) of a job on job choice.

5. Desire and need for money and security have been found influential.

6. Urban-rural differentials in interests have been found, but their nature is not consistent from study to study.

7. Varying, but probably lesser, degrees of importance have been found for: school, work experience, friends, guidance, working conditions, desire for freedom, and for adventure or excitement or glamour. There is little evidence that subjects considered their own abilities.

8. The importance of these factors may vary with the population studied. For example, teacher groups show more reaction to their own teachers and to a desire to work with people.

9. Subjects often do not mention influences for which there is other fairly clear evidence.

10. The relationship of interest to personality is gaining increasing attention, but little definitive work has appeared. In general, "job personalities" have not been successfully distinguished with currently-available measures of personality, although some scattered reports of success encourage further study in this area.

11. Results sometimes yield little of value except their indication of the need for precise work in studying what may be behind an interest. For example, the finding of "interest in it" as a leading influence on vocational preference (interest) appears to indicate what could be called "circular research."

Such inconclusive results are indicative of the need for improved research methodology, and review of reported studies does indicate methodological weaknesses that could account for much of the disappointing results.

2. *Evaluation of Methods Used in Studies of Interests*

a. *The major methods and techniques.* There are three major methods used for research on vocational interests: the cross-sectional study, the longitudinal study, and the case history. Aside from these, there are three major techniques used to collect data pertinent to the subject: the questionnaire, the test or inventory, and the interview.

The cross-sectional method involves getting data from a group of subjects of different ages. The data are collected over a relatively short period of time. This method has been used most extensively to date in vocational preference research. It has the disadvantage that inferences as to vocational interest *growth* which may be drawn from the data are deduced by comparing responses of different subjects. Usually the data include only the current preferences of the subjects, and only one preference at that. Those studies which ask for what the interests actually *are*, have not succeeded in getting an adequate picture of the *development* of the interests of the groups they have studied. Those that ask for the most important reasons for interest have not been at all conclusive.

The truly longitudinal study in this field would follow a group of subjects from early childhood into adulthood, with periodic checks on the current vocational preferences and the reasons for them. Control groups would be needed to gauge the effect of the questioning itself. This sort of study apparently has not been made, perhaps because of the span of years which it would require. Yet the follow-up of interests is of a longitudinal nature, and provides some opportunity to study changes of interest with age; however, it has primarily been used to study the frequency of change, not the nature thereof.

The case history method is in a way a combination of the two methods just described. It is cross-sectional in that its data is collected at one time; it is longitudinal in orientation, seeking information pertinent to the entire developmental period under consideration. The case history has the disadvantage of relying frequently upon recall of past events; however, recent studies have shown that recall of childhood memories and work histories may reasonably be used in study of groups (19, 1948; 9, 1950). This method has the advantage that it is aimed at getting as much as possible of the pertinent data relating to the past history of the matter under consideration.

Of the three techniques used for collecting data, the questionnaire is by far the most frequently used. It has the decided advantage of economy in administration. It has the disadvantage of limiting the data to what the subject can both perceive and verbalize at the particular moment—there is no possibility of rephrasing a question or probing for more relevant answers. The very ease of administration appears to have appealed to the careless student, who has used imprecise wording and undefined terms.

The scores of interest tests or inventories do not of themselves yield data directly relevant to the development of interests. They have been used, in both cross-sectional and longitudinal studies, to study changes in interest within a group, and to compare groups of different ages. However, most of the inventories are limited to use with high school subjects or older people, and thus are of no value in tracing the waxing and waning of early interests.

The interview has several disadvantages: it requires much more time and effort to administer, and it poses a problem in recording the data—simultaneous recording is necessary for accurate records. It is the most difficult of three techniques to standardize in administration. On the other hand, the interview allows for probing where necessary, for picking up misunderstandings, for rephrasing questions. It permits use of less structured questions. Yet a standard interview schedule yields data comparable from subject to subject. If done carefully, it can be more thorough; it permits

the inclusion of less verbal subjects—those who could not or would be disinclined to answer a written questionnaire.

b. Weaknesses in research method. In addition to those weaknesses just pointed out with reference to specific methods and techniques, the following situations are frequently met:

1. Much research appears to be rather superficial because of: (a) Poor phrasing of questions. For example, Orten (14, 1948) asked "What experiences have you had which either caused you to want to become a teacher or contributed to your fitness to become a teacher?" His question is "double-barreled" and confuses the issue; yet he entitles the report of his data "Why Do They Want to Teach?" (b) No standardization of questions. When the investigators report just what questions they asked, it is apparent that different results could be expected from different studies. The work of Gilger (5, 1942), Trow (18, 1941), Jones (8, 1940) and Williamson (20, 1939) give support to the need for very careful wording. Strong (15, 1943) argues that respondents confuse "work planning to do" with "a specific job offered and accepted." (c) No standardization of the situation in which questions are asked. Strong (15, 1943, p. 29) says "If students are required to express a choice (as at registration), practically all will oblige the authorities. If there is nothing at stake, a considerable minority will give no choice." (d) No standardization of classification or categorization of jobs. Results can be changed decidedly by a change in classification: a study which fails to combine the several sorts of engineer (mechanical, civil, etc.) may omit engineering from the most popular choices, where it would fall if these were combined. Similarly, Menger (13, 1932) lists stenographer and secretary separately, although it might be questioned how many third grade children would know there was a difference. (e) No clear definition or categorization of influences on vocational interest. For example, the categories "family," "parents," "elders," and "relatives and friends," if grouped differently, yield quite different results. (f) No out of school population. All of the studies reviewed were based on school students, ranging from third grade to college graduate students. At the lower levels, this might not bias conclusions, but as the population becomes more select at the higher levels, extreme caution should be used in generalizations. (g) Use of emotionally toned words. A rating scale seeking the importance of status (3, 1948) weighted (frowned upon by some people" as indicating more status than "considered an average job." Such judgments, which to this writer lack even face validity, should be supported by research before they are used to investigate interests. (h) The attributing of cause to con-

current factors when only co-existence has been proved. (i) No checks of the recording methods of those studies done by interview. There is no reported study in which the biases or errors of the interviewer were investigated.

2. No study has been made since Fryer in 1931 (11) of the waxing and waning of interests.

3. Much of the work on influences on vocational interest depends on the subject's insight into his own vocational development. There is evidence that such insight is inadequate.

4. There has been almost no use of statistical tests of hypotheses or of differences found.

Thus much of the research on interest development has apparently been very superficial. The work reveals weaknesses in technique that are at least somewhat avoidable, and it shows failure to approach some of the basic problems directly. One of the outstanding gaps is the lack of data that trace the developmental history of the individual's interests. Another outstanding methodological need is for an objective approach to the study of influences on interests. Subjects apparently do not always have insight into their own vocational choices, so that methods must be used that will go beyond their expressed reasons as the source of data. There is vital need for a better method to estimate the importance of influences: it must not merely study the correlates of interests, but should include estimates of influence which can be confirmed by observer agreement, and tested statistically. There has been little work with older subjects, work which could bridge the gap into the early adult years.

Insofar as findings are concerned, then, there have been no outstandingly definitive results from the studies to date; in fact, there has been decided disagreement. The large amount of speculation that appears in a survey of current knowledge of influences on vocational choice is evidence of the inconsistency and generally unsatisfactory nature of the data. Review of the research reveals so much that we need to know that no one study can hope to fill all of the gaps. This study sought to avoid the weaknesses of methodology described, and sought to assemble data which would fill some of the gaps in the basic information about how vocational interests and choices develop.

3. *The Plan and Procedure of the Present Study*

This was primarily a descriptive study, aimed at extending knowledge of the unfolding of interest patterns into the adult years. It sought to uncover age trends in both interest development and in the patterns of early

work careers. That is, it investigated when vocational preferences first appear, what they are when they do come, how they drop off or change. It aimed to answer these same questions about early work careers. Then, finally, as will be reported in a second paper, it sought a more comprehensive answer to the question "why" than has been thus far obtained by asking subjects directly or by studying correlated factors; this it did by drawing inferences by global judgment of the total interview record.

a. *The population studied.* A total of 193 subjects were interviewed: 75 women teachers, 75 male teachers and 43 factory workers. The teachers were selected to represent a professional type of work, and to permit study of sex differences. Inclusion of the factory workers made possible study of contrasts between professional and non-professional groups, and also involved some of the non-school, less verbally facile subjects who have tended to be overlooked in vocational interest research. The number of subjects was limited by both the time required by the interview and the relative inaccessibility of subjects.

All of the subjects selected were between the ages of 28 and 40, so that they would be somewhat settled in their field of work, would have met such economic and social factors as the depression and World War II at somewhat the same age, and would not be too far removed from the childhood ages they were to recall.

Although volunteer samples have been strongly criticized (17, 1948), it is impossible to get data on interests from people who will not coöperate. The teacher subjects represented nearly the entire sample of those teachers in the 27 high schools visited in central New York State who did not teach art or music, were of the proper age, and had free periods on the day the interviewer visited that school. The factory workers were obtained through the personnel departments of three manufacturing concerns in Syracuse. It is not known how representative they were of workers in this community; they were all that could be contacted, at home or in the factory, in the time available.

The mean age of all three groups was 34. There were some differences: the workers had less education than the teachers, although 48 per cent of the former were high school graduates. The factory workers' nationality backgrounds were weighted towards Italian, the women teachers' towards English. Although the teachers' parents had had more education, and more often were small proprietors than the factory workers' parents, there was no great evidence that the groups came from different "classes."

b. *The procedure in collecting data.* The most appropriate method to

get the desired developmental data, in the time available, appeared to be the case history, obtained by interview, so as to avoid the superficiality of earlier studies.² The interview form was constructed over a period of six months, with changes made after trial use. Open-ended questions, similar to those developed by Likert (11, 1947) seemed especially appropriate to elicit the recall of job preferences and early work histories. The subjects reported "the very first thing you remember wanting to do or be," and then the next. Appropriate standard probing questions were included to discover the ages when each interest or job was held, and reasons for each. Then the subjects reported on their "very first job for which you were paid," and on what they liked and disliked about this and later work. The final portion of the questionnaire was more highly structured, and included specific questions relating to the extent which family, school, etc., tended to influence the development of the subjects' thinking. The writer did all of the interviews, recording each subject's response directly onto the interview schedule. Electrical recordings were made of one-fourth of the interviews for a check on the accuracy of the records. The written "vocational case histories" provided the basis for the tabulations used in the study, and the record which was judged for the importance of different influences on interest growth.

c. Method of analysis of the data. The objective of this study was to find the picture of the development of vocational preference and actual work choice of the subjects. This was accomplished by tabulations of the interests and jobs started and held at each age by the subjects. Comparisons were made throughout the study between the women and men teachers, and between the male teachers and the male factory workers. For the most part, comparisons were accomplished by estimating the significance of the differences of percentages, using 148 and 116 degrees of freedom for the 150 teachers, and 118 male subjects (see Johnson, 7, 1949, and McNemar, 12, 1949, p. 75). For some over-all estimates of differences of the groups, the chi square test was used.

d. Reliability of the data. Since the subjects were not available, a check on the data by re-interview was not possible. It was felt that the interview technique used was as adequate as that of other interview studies, and that the present study, restricted to recall of a specific area (vocational interest) and using probing questions, could be expected to yield data as dependable

²This decision was made before Ginzberg, *et al.* (6, 1950, p. 176) made their statement relative to their choice of the interview method.

as that of Kinsey (10, 1948), or Keating, Paterson and Stone (9, 1950). Such data appeared valid for use with group comparisons.³

One approach to reliability estimation is to study the stabilization of the results—that is, the reliability of the general findings. If one-half of the sample yield the same sorts of data and generalizations as the other half, it is reasonable to assume that added cases would not change further the generalizations. Therefore, 24 comparisons were made between the even and odd-numbered subjects of each group. Only four indicated that the half-groups differed (p of 0.01) or approached a significant difference (p of 0.02-0.10). Thus, in general, the two halves of each group yielded data that were not reliably different.

A check was also made on the reliability of the tabulations made from the written interview records. If the interviews had been perfect, with the ages stated at every turn, a second tabulator, listening to an audible record, should have produced exactly the same tabulations. Two graduate students in psychology recorded and tabulated the interests and jobs mentioned by 21 subjects. These independent raters agreed with the writer's tabulations on 78 of 86 interests mentioned, or 90.6 per cent; agreement on jobs mentioned was 94.3 per cent. In tabulating the "age of beginning" and "age of ending" of 78 interests and 131 jobs, the independent workers agreed with the writer (within 2 years) 85.3 and 89.3 per cent of the time. Their discrepancies were due partly to the poor recordings, and partly to interpretation of phrases relevant to time, as "two or three years." Errors such as those found of 10- or 11-year differences cannot be adequately defended, but might be found in many interview studies had the authors recorded and checked their tabulations. Nevertheless, the consistency of age tabulations appeared to be sufficiently high that understandings might be drawn from the age data for an over-all picture.

B. THE DEVELOPING PATTERNS OF VOCATIONAL INTEREST AND JOB CHOICE

1. *Emerging Vocational Preferences*

a. *Ages when interests appeared.* The ages at which the subjects first had any interests at all are shown in Table 1. The most striking difference

³An interesting slant on the problem of accuracy of recall was reported in a personal communication by Charles Langmuir. He asked adults about recall of vocational interests of college students, and frequently got the reply that to expect accurate recall of such things was preposterous. He then asked the individual about his own interest growth, with the frequent response: "Well, in my case it's different, because when I was . . ." (followed by a specific vocational memory related to a specific age or to some event which was vividly and reliably remembered).

TABLE 1
PER CENT STARTING FIRST VOCATIONAL INTEREST AT EACH AGE

	Age																			
	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	27	
Women Teachers	1	5	11	12	8	12	7	17	3	9	7	3	4	1						
Male Teachers	3	3	12	9	11	11	1	8	5	13	3	5	8	3	4	0	0	1		
Factory Workers	2	0	2	5	0	5	2	2	5	16	12	12	7	7	7	2	2	0	2	

is that the factory workers developed interests later. Their mean age of first interest was 12.9 compared with 9.7 for the male teachers and 8.8 for the women teachers.⁴ The latter, although not significantly earlier with respect to mean age, did surpass the male teachers at ages 4, 6, 8, 9, and 10. Thus interest in vocation seemed to have made its impression earlier in those who eventually became teachers; the non-professional group lagged considerably behind.

Figure 1 goes beyond the first interests to show the picture of how inter-

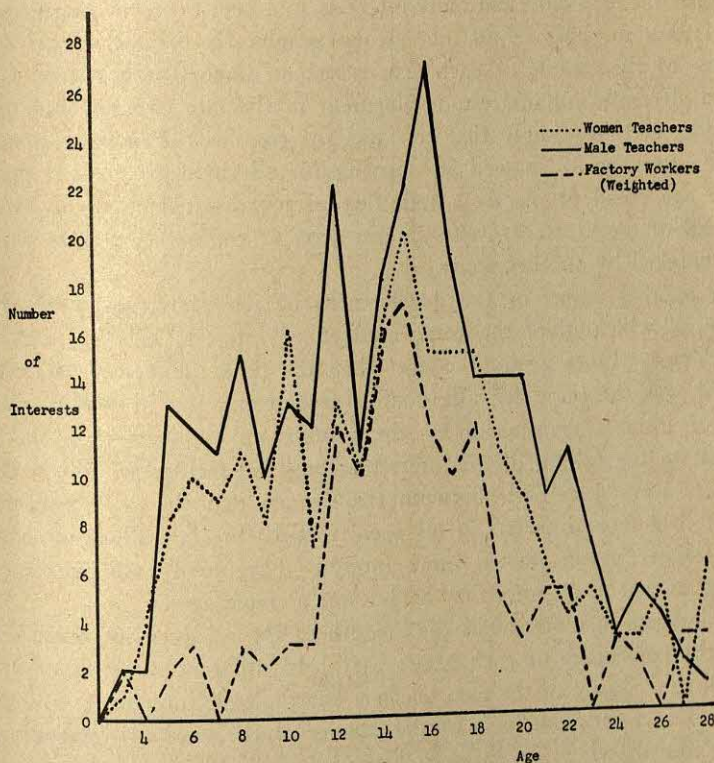


FIGURE 1

NUMBER OF VOCATIONAL INTERESTS BEGUN AT EACH AGE

The number reported begun by 75 women teachers, 75 male teachers, and 43 factory workers. The latter's results were weighted by 1.744 so as to be comparable to a sample of 75.

⁴CR ("t") between male and female teachers was 1.5, not significant at the 10 per cent level of confidence. Between male teachers and factory workers, the "t" was 4.1, significant at the 1 per cent level of confidence. In neither case did the variance ratios (F) reach significance at the 10 per cent level of confidence.

ests continue to appear and multiply during the early life-span. Specifically, it shows the number of interests *begun* by each group at each age. Reference to the ages at which the groups reached a peak in starting new interests, and to the age at which they first started half of the peak number, reveals that the men and women teachers followed much the same pattern, whereas the workers lagged. The fact that the number of interests begun by the men teachers was greater at most ages, is not inconsistent with Table 1 nor with findings reported in earlier studies; later data will show that the men both had more interests, and kept them for shorter periods.

The data for ages 29-40, not shown graphically because of the decreasing size of the sample at each age, raised an important question: Is there a "recrudescence" of interest development in the late twenties and thirties? The data certainly imply that this was the case, as the number of interests started after age 27 showed an upswing for all three groups. If this were a true age trend or merely a reflection of job dissatisfaction, or the result of recall of more recent events on the part of the older subjects will have to be checked by another study.

Yet another aspect of the development of job preference is revealed in Figure 2, which gives the total number of interests held by each group at each age. Once again, the age when the peak number of interests was reached was the same for all groups (seventeen), but in general the high level of interest maintained by the professional subjects stood in strong contrast to the falling off of interest among the factory workers. Slighter contrasts were discernible between the women and the male teachers; the former held more interests at all ages from 10 to 16. Thus, although the male teacher group started more interests (Figure 1), theirs were more fleeting so that as a group they held fewer at most ages.

To indicate in more detail the breadth of the subjects' interests, Table 2 gives the percentage of each group that held different numbers of interests at each age. More of the men teachers had multiple interests at all ages but age 21. Dramatic emphasis of the paucity of the workers' interests is seen in the row showing "no interests," as well as in Figure 3. At no age did more than 67 per cent of this group have any interest; at only 10 age levels did more than half of them have any interest at all. The percentage of workers expressing any interest was reliably lower than the percentage of male teachers, throughout the age span from age five.

Comparison of the first age when each entire group had at least one interest shows that the women teachers were earliest (age 18). This group surpassed the male teachers in percentage with at least one interest, at every

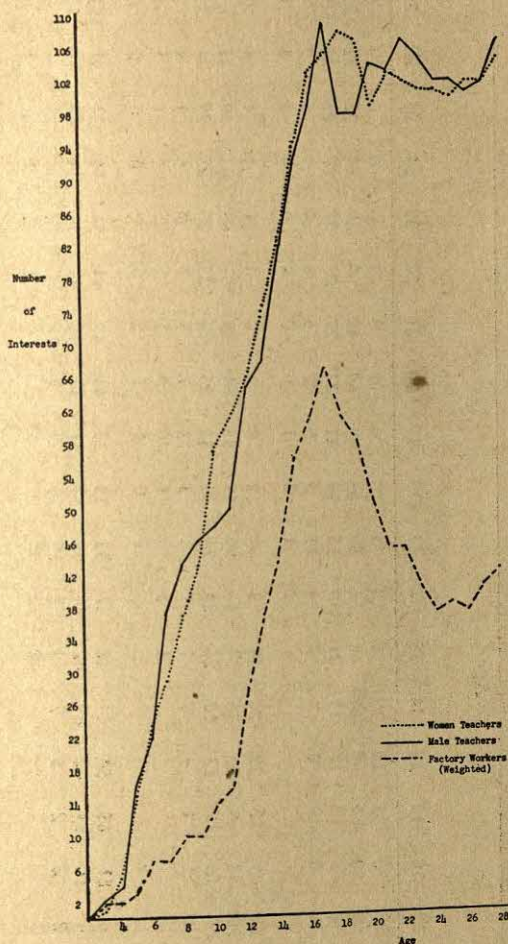


FIGURE 2

NUMBER OF VOCATIONAL INTERESTS HELD AT EACH AGE

The number reported held by 75 women teachers, 75 male teachers, and 43 factory workers. The latter's results were weighted by 1.744 so as to be comparable to a sample of 75.

age except age three. However, this difference was statistically significant only at ages 10 and 11. This fitted in with earlier studies which showed that girls report interest more frequently at earlier ages, but the data here made clear that one must be very careful how statements about the fact are worded. Although the women more frequently had at least one interest, at some ages the male teachers actually had more interests in the aggregate because of

TABLE 2
PER CENT OF SUBJECTS HAVING DIFFERENT NUMBERS OF VOCATIONAL INTERESTS AT EACH AGE

	Number of in- terests	Age																										
		3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	
Women Teachers	0	99	93	80	72	64	52	47	27	32	23	16	12	5	3	3	0	0	1	0	0	0	0	0	0	0	0	
	1	1	7	20	28	35	47	48	68	60	65	67	67	67	68	69	68	68	69	68	69	69	72	72	71	68		
	2				1	1	1	5	4	8	12	17	17	21	19	13	19	21	23	27	25	27	27	23	23	24	24	
	3								1	0	0	0	4	7	9	13	9	9	7	5	5	4	4	5	3	4	8	
	4														1	1	3	1	0	0	0	0	0	0	3	1		
Male Teachers	0	97	95	81	76	64	55	56	51	49	32	28	20	11	8	8	8	8	3	0	0	1	3	4	4	3	3	
	1	3	5	13	15	25	35	32	36	33	49	55	55	56	53	59	57	55	64	69	67	65	63	59	63	65	64	
	2			5	9	9	8	9	12	16	17	13	19	25	29	24	31	33	25	24	25	25	28	29	24	23	24	
	3					1	3	3	1	1	1	4	7	7	7	8	3	3	7	5	7	7	5	7	8	8	7	
	4													1	3	0	1	0	0	1	1	1	1	1	1	1	1	
Factory Workers	0	98	98	95	91	91	86	86	84	81	65	54	44	42	37	33	33	35	44	49	44	54	54	54	54	51	49	
	1	2	2	5	9	9	14	14	14	16	33	42	49	47	54	49	54	44	47	49	51	54	44	44	44	47	49	
	2								2	2	2	5	7	9	7	19	14	9	9	12	7	2	0	0	0	0	0	
	3													2	2	0	0	0	0	0	2	2	2	2	2	2	2	
	4															0	0	0	0	0	2	2	2	2	2	2	2	

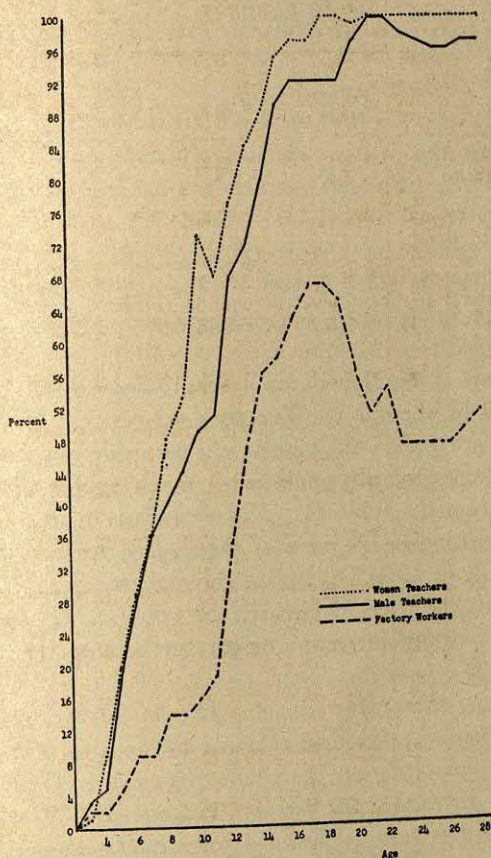


FIGURE 3
PERCENTAGE REPORTING ANY VOCATIONAL INTEREST AT ALL AT EACH AGE

the multiple interests of those who had any. Those male teachers who had interests had more of them (Table 2), so that at most ages the smaller percentage of them (Figure 3) held almost as many or more interests than the women teachers (Figure 2).

Table 3 indicates once again that the teachers started interests earlier, and that the factory workers held the fewest interests, whereas the male teachers surpassed the women teachers in number of interests held.

Thus the pattern of development of interests, as reported by the subjects, can be described: first interests developed at a wide range of ages (3 to 27), with the professional subjects centering around age 9 to 10, the factory

TABLE 3
FINDINGS BEARING ON DEVELOPMENT OF VOCATIONAL PREFERENCES

	75 Women teachers	CR	75 Male teachers	CR	43 Factory workers
Mean age of first recalled preference or interest	8.8	1.5	9.7	4.1**	12.9
Mean number of vocational interests in life to date*	3.3	4.1**	4.3	5.7**	2.4

*Includes those interests begun at ages 29-40 for those subjects older than 28.

**Significant at the 1 per cent level of confidence; the variance ratios (F) were not significant at the 10 per cent level of confidence.

workers around age 13. New interests continued to start throughout the age range covered by the study. The peak was at age 15 to 16, but there was pointed evidence for an upswing in the late twenties and early thirties. The factory workers generally indicated a much poorer vocational orientation than the professional subjects. The complexity of the pattern of interest growth has been indicated: *more* of the women teachers held interests at most ages, especially at 10 and 11. But the male teachers as a group *started* more interests. In the over-all picture, the two teacher groups *held* about the same total number of interests at each age because the fewer men held more dual interests.

b. How interests waxed and waned. The first interests of the male subject groups were more varied than those of the women. With 49 per cent giving teaching as first interest, the women listed only 17 fields of interest; the male teachers listed 34, the factory workers, 24. The length of time these interests lasted is seen in Table 4; clearly the women held their first interests longer, as would be expected with half of them giving teaching first.

TABLE 4
THE NUMBER OF YEARS VOCATIONAL INTERESTS WERE HELD

	75 Women teachers	CR	75 Male teachers	CR	43 Factory workers
Mean number of years <i>first</i> vocational interests lasted	14.8	4.8**	7.8		8.0
Mean number of years <i>all</i> vocational interests lasted	10.7	***	8.5	***	6.7

**Significant at the 1 per cent level of confidence; however, the variance ratio (F) was significant at the 2 per cent level of confidence.

***Means derived from tabulations which did not permit computation of SD 's or CR 's.

A check of *all* interests reported revealed that the "prestige" positions stood out among the men teachers, but the factory workers did not remember those interests as clearly. An emphasis appeared on "business"; for the teachers, this included work with industry and foreign companies, while for the workers it was exclusively a business of their own. This business, and professional ball-playing and engineering, were the only fields mentioned by more than one-tenth of the group. In contrast, 11 fields besides teaching were recalled by more than one-tenth of the men teachers; five, by more than one-tenth of the women teachers. This indication of the greater homogeneity of the women's interests was supported by the fact that more of the men held multiple interests, as well as by the women's concentration of first interest in teaching, already pointed out. However, both teacher groups reported 27 per cent of their interests in the five most popular fields; 33 per cent of the workers' interests were in *their* five most popular areas.

A breakdown of the interests as to the ages when they were held is revelatory of age trends; it is presented in Tables 5, 6, and 7. There can be studied the fields that hit peaks of popularity, that dropped out early, or that developed late in comparison to others. The male teachers especially showed pronounced patterns: the childhood fancies came and went—fireman and railroad engineer; most fields other than teaching dropped out by age 28. The trends among the factory workers were less clear, as might be expected in view of their smaller sample and general paucity of interests.

c. Age trends in reasons for interests. No clearcut trends that bear graphical or tabular treatment were apparent, but certain changes were noted. For the women teachers, "family" remained near top frequency until after 19; "liking for school subjects" rose slowly in relative importance throughout the life span studied. Direct influence of school and teachers increased in importance to age 19, then dropped off. "Experience in related activity" first appeared at ages 10-14, and became relatively more important with the passing years. Finances and need for immediate income made themselves felt more after age 15, and friends were of more influence in the twenties than earlier.

For the male teachers, too, there was some drop in influence of the family after 19, whereas the influence of school and teachers was relatively important for them through the mid-twenties. Experience, and security and money, became increasingly important from age 10 on, and consideration of ability or training appeared after age 15.

The factory workers referred only to the influence of school and family, and then reported these gave way to the effect of "similar or related experience," friends and money.

TABLE 5
PER CENT OF WOMEN TEACHERS HOLDING CERTAIN INTERESTS AT EACH AGE

	Age																											
	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28		
Lawyer						1		1	1	3	3	4	5	4	4	4	3	1										
Nurse		1	3	5	5	3	1	7	5	7	9	6	4	4	1	3	1			1	1					1		
Teacher		4	8	16	20	27	35	43	52	51	55	61	72	77	79	87	92	93	97	99	97	96	96	96	96	96		
Artist	1	1	1	1	1	1	1	1	1	3	3	1	4	4	1	1	1											
Actress, musician			1	1	3	5	7	8	8	9	9	9	9	8	5	5	4	1	1	1	1	1	1	1	1	1		
Writer, journalist				1	1	1	3	3	1	3	3	4	4	5	4	3	5	4	4	4	4	4	4	4	4	5		
Fashion designer						1	1	1	1	3	3	3	4	4	4	3												
Librarian									1	1	1	4	5	5	5	7	7	5	7	7	9	8	7	7	8	8		
Doctor, surgeon										1	1					1	3	4	3	3	1	1	1	1	1	1		
Social worker												1	3	3	5	4	4	4	4	3	3	3	3	3	3	3		
Guidance and counseling																												
Other professions								1	1	1	1	5	7	12	15	11	9	9	8	8	8	9	11	13	12	12		
Waitress			1	1	1	1	1	1																				
Office work											3	4	4	7	9	9	7	3	4	5	4	5	5	5	5	3		
Store clerk, merchandising					1	3	3	3	1	1	1	1																
Homemaking, cook								3	4	4	4	4	3	3	3	3	3	3	3	3	1	1	1	1	1	1		
Own business															1	1			1	1	1	3	1					
Armed services																			1							1		
School administration																			1									
College teacher																					1	1				1		

TABLE 6
PER CENT OF MALE TEACHERS HOLDING CERTAIN INTERESTS AT EACH AGE

Field of interest	Per cent holding at each age																											
	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28		
Service (delivery)	3	4	7	3	3	1	1	1	1																			
Minister, Priest		1	1	1	1	4	4	4	4	4	3	3	3	3	3	3	3	3	3	1	1	1	1	1	1	1	1	
Cowboy			1		1	3	4	3	1	1	1																	
Policeman			1	1	5	5	4	3	1	1																		
Farmer			1	3	3	1	3	3	4	4	4	4	3	1	3	3	4	4	1	1	3	3	3	4	4	4	4	
Trade (mechanic, plumber, etc.)			1	3	3	3	5	5	5	7	4	4	5	8	4	5	5	5	5	5	5	5	5	5	5	5	5	
Engineer			1		1	3	1	3	3	4	4	5	8	8	8	8	3	4	3	1	1	1	1	1	1	1	1	
Lawyer			1	1	1	1	1	3	3	5	7	9	8	8	7	8	8	8	7	4	3	1	3	3	3	3	3	
Actor, musician, dancer			1	4	4	4	4	4	7	5	5	7	9	9	7	3	1	1	1	1	1							
Fireman			3	4	9	8	7	4	4	3	1	1	1	1														
Other professions				1	1	1	1	3	5	7	9	12	15	13	16	12	15	11	9	7	4	4	5	5	4	4		
Railroad engineer				5	12	12	11	8	4	1																		
Aviator						3	4	4	1	3	3	4	4	4	4	4	5	3	4	4	4	4	4	1	1	1	1	
Doctor						1	1	3	3	13	15	19	21	20	13	9	9	7	3	5	7	4	3	3	1	1	1	
Pro ball player						3	4	7	9	11	11	12	9	5	4	4	4	3	3	3	3	3	1	1	1	1	1	
Teacher								1	1	4	12	17	21	31	43	52	56	72	81	89	89	92	92	92	95	96		
Own business								1	1	1	1	1	1	1	1	1	1	1	1	1	1	3	3	4	4	4		
Writer									1	3	3	3	4	4	3	3	3	3	3	1		3	3	3	3	3	3	
Salesman															1	3	1				1							
Business, banking														3	5	7	8	9	8	9	9	5	9	5	5	7		
Counseling and guidance																										3		
School administrator																								1	1	1		
College teacher																										1		

JOSEPH L. NORTON

TABLE 7
PER CENT OF FACTORY WORKERS HOLDING CERTAIN INTERESTS AT EACH AGE

	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Cowboy			2	5	5	5	2	2																		
Policeman									2	2	2															
Fireman			2	5	5	5	2	2	2	7	7	5	5	5	5	2				2	2					
Railroad engineer										5	5	5	5	5	5	2										
Pro ball player						2	5	5	5	7	7	12	9	9	12	9	5	5	2	2	2					
Musician, dancer	2	2	2	5	5	5	5	5	5	5	5	5	5	2	2	2	2	2	2	2	2	2	2	2	2	2
Aviator						2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Undertaker											2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Minister, Priest										2	2	2	5	5	2	2	2	2								
Engineer										2	9	7	5	7	7	5	5	5	7	5	5	2	2	2	2	2
Doctor												2	5	2	2											
Teacher															2	2	2									
Coach														2	2	2										
Other professions													2		5	7	9	7	2							
Own business								2	2	5	5	5	12	14	12	9	12	9	9	12	9	9	9	9	9	9
Salesman															2	2	5	5	5	7	7	9	9	9	9	12
Store clerk																									2	5
Mechanic											2	2	2	2	2	2	2	2	2	2	2					
A trade (printer, tool maker, carpenter, plumber)									2	2	5	9	7	9	7	9	9	7	5	5	5	7	7	7	7	5
									2	2	2	9	9	12	14	14	12	12	9	9	9	9	9	9	7	7

All of these reasons did appear to reveal an expansion of the subjects' world from the close-knit home environment to the broader horizons of school, and on to involvement with the forces of society. The realization of the need for money and security, and the awareness of interest in people, came along in the older ages included in the study.

2. *Emerging Vocational Experience*

a. *Part-time and summer work.* The analysis of the subjects' work experience was complicated by the intrinsically variable nature of work experience. Jobs may be part-time throughout the school months or the entire year, full-time for the summer only, or full-time year-in-year-out. Some subjects might have been paid in their own homes for household chores; others may not have been paid at all for years of hard work at home on the farm. In view of the spasmodic nature of the experiences, both part-time and summer jobs have been considered together throughout this study. Work done at home has been excluded from all reports of data because it was so difficult to distinguish between real work and an occasional chore done for an allowance.

Figure 4 reveals that the male teachers were ahead of the factory workers and the women teachers in getting part-time jobs, and that the workers dropped from their peak number at an earlier age. And Figure 5 shows that the male teachers out-stripped the factory workers in the number of jobs held, as well as in early start. Even the women teachers, although they lagged in the early 'teens, eventually surpassed the factory workers. The latter's pattern differed considerably from the teachers' ages of peak part-time work load. The somewhat surprising fact that the male teachers seemed to have started work earlier than the factory workers was confirmed by the mean age of *first* work experience of any sort (part- or full-time) as shown in Table 8. This may indicate that the professional group was made aware of the ultimate need to earn a living earlier than the factory workers, so that they also became more aware of the need for preparing for a life work. It may be that the factory workers' parents wanted to keep their children from appearing to "have to work," and so actually did them a disservice by failing to stimulate them to think vocationally. This implication should be carefully investigated. It seemed clear that the male teachers were not from more deprived homes, nor was there evidence to support any subjective estimate of great "class" or "level" differences between the male groups; certainly their backgrounds overlapped. Yet the teachers clearly revealed either more initiative or more push from their families in starting some remunerative work prior to age 13. This point

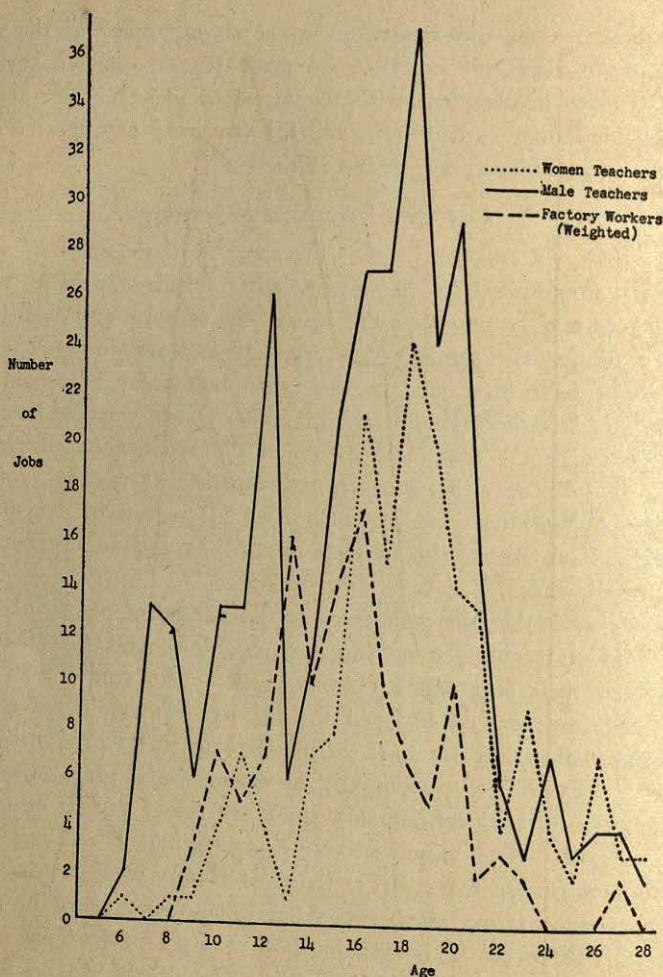


FIGURE 4

NUMBER OF PART-TIME OR SUMMER JOBS BEGUN AT EACH AGE

The number begun by 75 women teachers, 75 male teachers, and 43 factory workers. The latter's results were weighted by 1.744 so as to be comparable to a sample of 75.

was supported by the fact that only 3 per cent of the male teachers' first work was in a full-time job, whereas 22 per cent of the factory workers' first jobs were full-time.

The women teachers, as might have been expected, showed a much slower start than the men, and had 27 per cent whose first work experience was full-time.

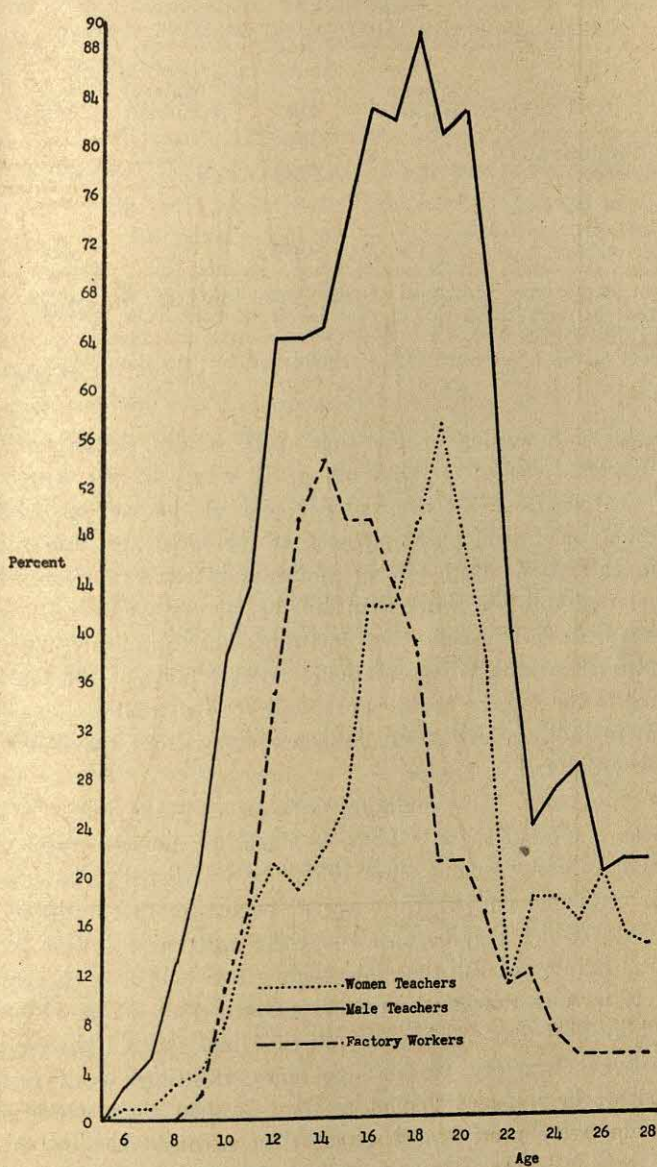


FIGURE 5
PERCENTAGE REPORTING ANY PART-TIME OR SUMMER JOB AT EACH AGE

TABLE 8
FINDINGS BEARING ON DEVELOPMENT OF WORK CAREERS

	75 Women teachers	CR	75 Male teachers	CR	43 Factory workers
Mean age of beginning <i>first</i> work experience	16.6	7.9*	12.0	2.7**	13.5
Mean number of job areas in which subjects reported working full-time	1.9	4.9**	3.0	5.9*	5.2

*Significant at the 1 per cent level of confidence; however, the variance ratio (F) was significant between the 2-10 per cent levels of confidence, and the data were not normally distributed.

**Significant at the 1 per cent level of confidence, but the data were not normally distributed.

The waxing and waning of part-time work is also shown in Figure 5. The male teachers reached the point at age 18 where 89 per cent held part-time jobs. In contrast, just over 50 per cent of the women teachers and factory workers ever held part-time jobs at the same age, and the latter's peak was much earlier. Both teacher groups maintained more part-time jobs after they started full-time work than did the factory workers.

There was little in the data about part-time jobs to indicate directly much effect on ultimate vocational orientation. Almost none of the jobs were in fields related to the subjects' adult work, and yet the reactions reported were not such as to indicate any strong influence away from the fields of work in which they did fall.

b. Full-time work. In tabulating data on full-time jobs, "a job" was used to mean a type of work. Thus, if a teacher moved several times but reported "I have been teaching since 1935," he was listed as having one job.

Whether compared on (a) their age of beginning first full-time job, on (b) the ages at which they began all of their different full-time jobs, or on (c) the total number of full-time jobs their group held at each age, the factory workers far out-stripped the teacher subjects (see Table 8). The percentage holding different numbers of full-time jobs showed the same results. Very clearly, the factory workers got into full-time work earlier; the professional workers lagged behind as they finished their education. The teacher groups were pretty much alike: they reached the peak number of full-time jobs held at age 25, and reached the half-way point at age 21 to 22. Of course, a few of the male teachers started full-time work at ages 14-17, while the women had none at these ages; also, the women forged ahead at age 21. Apparently the latter finished their education earlier, per-

haps because they showed more singleness of purpose, as revealed above. The male teachers as a group were more variable in their age of beginning full-time work than either other group.

It was noticeable that more of the factory workers switched "job area" in a single year, and reported holding more different types of work in the aggregate (Table 8). In the same vein, 55 per cent of the women teachers and 17 per cent of the male teachers had held only one type of work (teaching), whereas only one factory worker (2 per cent) had remained at the same type of work throughout his early adult years. Thus the earlier full-time work, coupled with their higher job turnover, support the finding in the section on interests, that this factory group showed a great lack of vocational orientation.

The women teachers led in starting their full-time work in their present type of work: 84 per cent of them started their full-time careers as teachers. A substantial number, 48 per cent, of the male teachers had teaching as their first full-time job; 26 per cent of the factory workers first started full-time work in a factory, although not necessarily in the same type of work. Only one factory worker reported military service as his first full-time work, but 14 per cent of the male teachers reported that.

There were no noticeable age trends in the types of full-time work other than the trend towards the current work and military service. The teachers worked in a variety of fields, with no emphasis on any field at any one age. The factory workers appeared to have been occupied in as many varied fields as the male teachers, with perhaps more emphasis on the laboring jobs. However, their earlier laboring jobs did decrease with age.

The subjects' reactions to their full-time work revealed no age trends, but they tellingly revealed the importance to their own feelings of others' opinions, in their responses to the question "What do you think about what other people think about this type of work?" There was no general consensus, even among the teachers, as 12 was the highest number of people to give the same answer. But the question did bring out from the teachers some indication of the prestige they felt attached or wished were attached to their profession. It was also evident that they do not agree with each other. There was less response from the factory workers.

Nearly twice as many men teachers as women teachers had at one time thought the work they were *then* in (not teaching) was the field in which they would plan to stay. Even more, although not significantly more, of the factory workers had once thought they would stay in another line of work. This may reflect their lack of vocational planning and foresight, and

their dependence on the employer for the decision as to their continuing in a field of work. The data also pointed out that only three-fourths of the women teachers planned to stay in teaching, whereas a few more of the male teachers expected to remain. In contrast, less than half of the factory workers thought they were in the line of work in which they would stay. Likewise, only one-fourth of the teachers felt that teaching was a second choice, whereas over half of the factory workers felt their work was second, third, or sixth preference.

C. SUMMARY

1. It was possible to get a picture of the rise and fall of a group's vocational interests, both in terms of specific vocational areas and total numbers of interests. All groups studied reached a peak frequency of interests begun, at age 15-16, although the professional and non-professional groups differed both in the number of interests held and the rapidity with which they developed the interests they did hold.
2. Vocational interests developed throughout the age span studied, through age 28; the fact a group had reached the early adult years did not cut off, although it reduced, the growth of new interests. The subjects' *first* remembered interests generally developed between ages three and 20; the mean age for the teachers was nine, four years ahead of the factory workers.
3. The male teachers held the most interests, the factory workers the fewest. Every comparison stressed the paucity of the factory workers' interest histories, their lack of vocational orientation. What interests they did have in the 'teens they appeared generally to have lost once they got jobs. The teachers maintained a higher number of interests throughout the age span studied. The *male* teachers held a higher total number of interests, but because they held them for shorter periods and tended to hold more multiple interests if they held any at all, their female counterparts had a higher percentage of the group who *had any* interest throughout most of the age span studied, especially at ages 10 and 11.
4. The nature of the women's interests was relatively stereotyped: 49 per cent had teaching as their first interest. The male teachers showed definite waxing and waning of the more common boyhood interests—cowboy, policeman, railroad engineer. But aside from teaching, medicine was their favorite. For the factory workers, professional ball player was the favorite; their other vocational interests spread over a proportionately broad range.
5. The subjects' current vocational goals did not consistently reflect

their earlier interests. Only 34, 65, and 44 per cent of the women teachers, male teachers, and factory workers felt they were progressing toward a vocational objective. However, each teacher had at least one vocational interest at the time of interview; only 70 per cent of the factory workers had one at that point.

6. There was some slight subjective evidence of an age trend in the reasons given for interests, from the immediate home environment to broader environmental, social reasons. The impact of the reality about the subjects was evident in reasons given for loss of interests.

7. The first work experience for all male subjects and for 75 per cent of the women started before age 20; some started as early as age six. The mean ages of first experience: women teachers, 16.6; male teachers, 12.0; factory workers, 13.5. The male teachers were clearly the earliest to start work; only 3 per cent of the group got their first pay for full-time work, whereas 27 and 22 per cent of the women teachers and factory workers did so. Fully one-fourth of the male teachers were paid for work at home, whereas only 14 per cent of the women and 9 per cent of the factory workers reported that stimulation towards thought of earning an income. For the men, there was little relation between the type of work done in their first jobs and their ultimate field of employment.

8. Their *part-time* work experience was the reason for the male teachers' showing earlier work histories. More of them held part-time jobs, and as a group they held more throughout the age span studied. The women teachers' part-time experience started more slowly, but hit its peak at the same age as the male teachers' peak number of part-time jobs—age 18 to 19. It was clear that the professional men were stimulated to think of remunerative work earlier than the non-professional men.

9. Part-time work experience was not much related to later fields of work. The only similarity was the contact with children afforded the women teachers by baby-sitting.

10. Entry into full-time work was earliest for the factory workers. While the two teacher groups did not differ in mean age of entry, at the ages 21 and 22 more of the women started work.

11. The women teachers appeared to be more stable in their full-time work; 55 per cent had had only one type of work—teaching. They had the fewest different fields of work, and the highest average time for staying in full-time jobs.

12. While their current fields (and military service) reflected the major

trends in types of full-time employment, all groups showed considerable work experience in other fields.

REFERENCES

1. BERDIE, R. F. Factors related to vocational interests. *Psychol. Bull.*, 1944, **41**, 137-157.
2. CARTER, H. D. Vocational interests and job orientation, a ten-year review. *Appl. Psychol. Monog.*, 1944, No. 2.
3. EDMISTON, R. W., & STARR, C. H. Youth's attitudes toward occupations. *Occupations*, 1948, **26**, 213-220.
4. FRYER, D. The Measurement of Interests in Relation to Human Adjustment. New York: Holt, 1931.
5. GILGER, G. A., JR. Declaration of vocational interest. *Occupations*, 1942, **20**, 276-279.
6. GINZBERG, E., GINSBURG, S. W., AXELRAD, S., & HERMA, J. L. The problem of occupational choice. *Amer. J. Orthopsychiat.*, 1950, **20**, 166-198.
7. JOHNSON, O. Statistical Methods in Research. New York: Prentice-Hall, 1949.
8. JONES, B. S. Relation of ability to preferred and probable occupation. *Educ. Adm. & Supervis.*, 1940, **26**, 220-226.
9. KEATING, E., PATTERSON, D. G., & STONE, C. H. Validity of work histories obtained by interview. *J. Appl. Psychol.*, 1950, **34**, 6-11.
10. KINSEY, A. C., POMEROY, W. B., & MARTIN, C. E. Sexual Behavior in the Human Male. Philadelphia: Saunders, 1948.
11. LIKERT, R. The sample interview survey. In Dennis, *Current Trends in Psychology*. Pittsburgh: Univ. Pittsburgh Press, 1947.
12. MCNEMAR, Q. Psychological Statistics. New York: Wiley, 1949.
13. MENDER, C. The significance of vocational choices of school children and college students. (Privately printed, New York, 1932.)
14. ORTON, D. A. Why do they want to teach?, *Phi Delta Kappan*, 1948, **29**, 343.
15. STRONG, E. K., JR. Vocational Interests of Men and Women. Stanford University, Calif.: Stanford Univ. Press, 1943.
16. SUPER, D. E. Vocational interests and vocational choice: present knowledge and future research in their relationships. *Ed. Psychol. Measmt.*, 1947, **7**, 375-383.
17. TERMAN, L. M. Kinsey's sexual behavior in the human male: some comments and criticisms. *Psychol. Bull.*, 1948, **45**, 443-459.
18. TROW, W. C. Fantasy and vocational choice. *Occupations*, 1941, **20**, 89-93.
19. WALDFOGEL, S. The frequency and affective character of childhood memories. *Psychol. Monog.*, 1948, **62**, No. 4.
20. WILLIAMSON, E. G. How to Counsel Students. New York: McGraw-Hill, 1939.

Graduate School
Alfred University
Alfred, New York

GENERAL MOTIVES AND INFLUENCES IN VOCATIONAL DEVELOPMENT*¹

Department of Psychology, Syracuse University

JOSEPH L. NORTON

A. THE PROBLEM AND THE METHODOLOGY USED

The review of the literature on development of vocational interests, given in the first article of this pair (3, 1953), prompted comment on the inadequacy of methods used in getting at the influences that are operative in vocational development. Too much dependence appeared to be given to the subjects' insight into their own interest growth, to inferences from the preferences stated by large groups of subjects, and to inferences from factors found to be co-existent with various interests. Therefore the purpose of this study was to try a new method of identifying both the general vocational motivation of each subject and those influences that seemed important in the development of his particular interests and work career. That is, a synthesis of each subject's vocational history was sought, a "clinical analysis" as Berdie (1, 1944) recommended. It involved obtaining global judgments of a series of complete vocational case histories—a judgment or rating as to the important motivations and influences operative in vocational development.

The data to which this method was applied, data described in more detail in the earlier article, were the "vocational case histories" of 193 subjects. These histories of vocational interests and actual job choices were obtained by interview with 75 women teachers, 75 male teachers, and 43 factory workers, aged 28 to 40.

The analysis was accomplished by two steps: the first was the categorization of the dominant drives and factors influencing the subjects. Two independent readers listed, for a series of case histories, those influences and drives that appeared in the series. Of a total of 77 items listed, there was identical agreement on 33, and comparison of the others revealed that they differed only in phraseology. A conference brought agreement on the wording, and the categories of drives and influences that appeared to be operative in the subjects were agreed upon. Six dominant drives or motives and 11 categories of influences were identified as evident in the histories of the subjects.

*Received in the Editorial Office on May 16, 1951.

¹Based on a portion of a dissertation submitted to Syracuse University in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The author gratefully acknowledges the helpful guidance received from Dr. Raymond G. Kuhlen.

The second step of this "synthesis" was to determine the relative importance of these factors to the subjects. Each interview record was rated (on a five-point scale) as to the importance of each of the drives and influences in that history. This method overcame the weakness of basing estimates of importance solely on the subjects' insight; it furnished the rater with a brief overview of the subject's whole developmental pattern, allowed him to integrate some of the minutiae into a coherent whole, and thus yielded a more complete picture of influences on interest growth. The method also permitted a check on observer agreement, as an independent rater reviewed and judged 50 randomly selected interviews for comparison with the first ratings. Tests of the relative importance of the categories were made by chi square. For example, in order to test whether a category was more important to the women or male teachers, the groups' distributions of high, average, and low ratings on the category were compared by the chi square test. Likewise, tests were made to estimate whether one category was more important than the others to the members of any one group; thus estimates were made of the relative importance of the categories, both from group to group and within each group.

B. THE FACTORS IDENTIFIED

In identifying the factors, operational definitions were sought—definitions built up from phrases which indicate clearly the meanings of the terms. Based on quotations from the very interviews to be rated, they were both more meaningful and easier to work with in the rating process. Some of the definitions arrived at were quite broad in scope, but each was rated as a whole. The definitions were as follows:

MOTIVES

1. *Status*

a. Vocational: have a sense of importance; have responsibilities; something not everyone can do; want to get into a profession; wanted to better self; I like the professional attitude.

b. Social: want prestige; it's a better class job; I feel I am above that; social approval is necessary; I want respect of others; have standing in the community.

c. Educational: I was taught to look down on non-college graduates; I always advise that an education is important even for ditch-digging.

2. *Association with People*

Like to work with people; don't like to work alone; like dealing with people; want to be liked by people; liked kids; want to be with people.

3. *A Satisfactory Living: Income and Security of Income*

a. Money: I needed an immediate income; wanted money so I could marry; for me money's the thing; I quit three times for jobs with more money; needed money to live.

b. Security: It's steady; dependable; the other job had no future; I need to consider family responsibilities in the future; a retirement plan.

4. *Helping Others*

Contribute to society; I can express moralistic values; want to help others; I can do more for people here.

5. *Self-Expression*

a. Creativity: like to build; like making things with hands.

b. Sense of accomplishment: like to have pride in work; want challenge.

c. Independence: want freedom; like to use own initiative; given own responsibilities; boss leaves alone.

6. *Variety*

Avoid monotony; like the excitement and glamor; disliked that because it was dull; it's the same thing over and over, so disliked; like it because it's different every day.

INFLUENCES

1. *Family: Father, Mother, Siblings, Brother-in-Law*

a. Direct advice, suggestion, decision.

b. Indirect influence, by example, moral teaching, insinuation.

c. Financial support for schooling or starting business.

d. Got the job for him or knew of opening.

2. *Other Relatives*

Other than above, as cousins, aunts, uncles. Same as for family.

3. *Friends*

Same as for family. Also, "went to school together."

4. *School (Education System)*

a. Teacher—same aspects as for family; also: example; liked teachers; got scholarship for me.

b. Guidance.

c. College within commuting distance.

d. College chosen offered only certain courses.

e. Scholarships only to certain college.

f. Particular school or teachers discouraged me so I quit early.

5. *Own Work Experience (Part- or Full-Time)*

- a. Earlier work in that line: liked, or disliked it enough to avoid that line.
- b. Had contact with that field while at work on something else, saw it and thought I would like it.

6. *Financial Need or Limitation*

- a. Couldn't afford training or schooling; financially unable to do what I wanted.
- b. Had to have any job I could get; nothing else available.
- c. Depression.

7. *Talent or Ability (As Viewed by the Subject)*

- a. Pick job to fit my abilities.
- b. Lack of talent: I am not talented enough for a concert musician; I couldn't aspire to such heights as the legitimate stage.

8. *Opportunity, Happenstance*

- a. Job was open in my home town and I took it; the factory hiring that day; I pounded the pavements, that day they needed somebody.

9. *Health*

- a. Into this sort of work: knee bad, needed warm place to work.
- b. Out of something else preferred: my knee keeps me from plumbing.

10. *Lack of Vocational Stimulation*

(I.e., high rating means that subject shows no sign of anything much to activate vocational interests.)

I never thought of it while in high school; I just drifted; I'm undecided always about what I want to do; never could make up my mind; never learned to look to the future; a fellow never settles down until he's 25.

11. *Early Extended Contact with that Sort of Thing*

(I.e., not influence of own work experience, but an extended contact with a vocational area through some other way.)

We lived next door to a print shop; my uncle a machine operator lived with us; I lived in an atmosphere of teachers; with my background I couldn't be anything else; as a child I loved to build things; I admired my best friend's father who was in that line of work.

It was obvious that the motives here defined must of themselves have causes. But if they themselves and the influencing factors could be identified and their relative importance judged, a step had been made in the right direction.

C. RESULTS OF THE RATINGS

To get an estimation of the importance of the six internal motivations and the 11 more external influences, each interest history was rated for evidence of the seventeen categories. The scale used for ratings was as follows:

- 4: Very dominant factor; possibly "sticks out all over."
- 3: Important factor to this individual.
- 2: Some influence noted.
- 1: Very slight indication of this factor; little influence.
- 0: Nothing in record to show that this factor entered in at all.

Raters were given the definitions, directed to read the interview record, and rate each one for all the drives and factors.

1. *Reliability of the Ratings*

To check the reliability of this method, after 50 interviews had been rated, a second rater read the same 50 records and rated them on the six motives and 11 influences. The correlations between the two sets of ratings for all but one factor were sufficiently high to permit confidence in group comparisons, although they were not as high as desired. The correlations indicated it was possible to approach observer agreement on the importance of all the categories studied except one—"talent."

The consistency of the ratings (intra-personal reliability) was investigated by correlating the writer's results of re-rating fifty interviews on all 17 categories after two month's time. These coefficients showed that the writer had, in general, better agreement with himself than with the other rater; also, they indicated that the agreement was not as high as desired in considering individual cases. All except "money and security," however, were of a level that indicated some stability in group ratings, and were useable for group comparisons. These correlations are shown in the first two columns of Table 1.

2. *Contrasts Between Groups*

Table 1 reports the average of the ratings given each motive and external influence, listed in order of the averages for the total group. The male teachers surpassed the women teachers on three motives and three influences, the factory workers on five motives and four influences. The workers showed the stronger influence only of the pressure of lack of money, and of a noticeable lack of vocational stimulation. Thus the male teachers showed more drive for money and security, but the factory workers evidenced more influence of actual limitation by finances. Clearly the male teachers received more of the higher ratings than did the other two groups. This supports the suggestion

TABLE 1

THE RELIABILITY AND THE RELATIVE IMPORTANCE OF THE MOTIVES AND INFLUENCES

The extent to which various motives and influences were revealed in the 193 vocational histories as indicated by mean ratings based on a five-point scale from 4 to 0. The marks > or < indicate a significant difference in the distribution of ratings (high, average, low) for the subject groups, as tested by chi square.

	Corre- lation between inde- pendent ratings	Corre- lation between repeated ratings	Mean for 193 Sub- jects	Mean for 75 women teachers	Mean for 75 male teachers	Mean for 43 factory workers
<i>Motivations</i>						
Self-Expression	.50	.53	2.14	2.15	< 2.48	> 1.53
Status	.49	.69	2.05	2.17	< 2.41	> 1.21
Association with People	.58	.71	1.95	2.15	< 2.37	> 0.86
Money and Security	.41	.35	1.93	1.52	< 2.31	> 2.00
Variety	.45	.50	1.38	1.35	1.36	1.53
Helping Others	.65	.88	1.03	1.11	1.41	> 0.21
<i>Influences</i>						
Family	.57	.77	2.55	2.88	2.72	> 1.67
School	.43	.71	1.89	2.19	2.33	> 0.61
Financial Need or Lack	.56	.80	1.83	1.39	< 1.92	< 2.51
Early Extended Contact	.41	.42	1.74	2.01	2.01	> 0.79
Friends	.56	.74	1.51	1.54	1.54	1.30
Own Work Experience	.35	.48	1.49	1.04	< 1.81	1.72
"Lack of Stimulation"	.47	.57	1.15	1.04	0.89	< 1.77
Talent (Consideration of)	.15	.46	0.83	0.96	0.88	> 0.51
Other Relatives	.67	.90	0.75	1.05	> 0.63	0.44
Opportunity	.42	.38	0.56	0.51	0.60	0.58
Health	.81	.90	0.34	0.13	< 0.35	0.70

that grew out of their earlier part-time work—that there were more and/or stronger forces acting on the male teachers as a group than on the others.

3. *Relative Importance of the Categories Within Groups*

Tables 2, 3, and 4 show which categories had significantly higher distributions of the ratings as tested by chi square. The categories are presented in order of their importance, as determined by the number of other categories each exceeded.² To supplement these data Table 5 reports the percentage of each group that received high ratings, and any ratings at all, on each category. The results emphasize that *many* influences operated to determine the subjects' interests. Indicative of such multiple influences are the fact that at least 25 per cent of each group were rated as indicating that each of the 6 motives had affected their vocational preference, and over 50 per cent of each group showed influence of 10 or more of the categories.

²This order closely approximated the order, for each group, of the average ratings.

TABLE 2
COMPARISON OF THE RELATIVE IMPORTANCE TO THE WOMEN TEACHERS
OF THE SEVENTEEN CATEGORIES

	Motives						Influences										
	1	2	3	4	5	6	1	2	3	4	5	6	7	8	9	10	11
Family (1)	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
<i>Association people</i> (2)			*	*	*	*	*	*	#	*	*	*	*	*	*	*	#
School (4)			*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
<i>Self-expression</i> (5)			*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
<i>Status</i> (1)			*	*	#	*		*	*	*	*	*	*	*	*	*	*
Early Ext. Contact (11)			*	*		*	*	%		*	%	*	*	*	*	*	*
Friends (3)							&			&		%	*	*	*	*	*
<i>Money and Security</i> (3)							%					#	*	*	*	*	*
Financial Need, Lack (6)												#	*	*	*	%	*
<i>Variety</i> (6)							&						*	*	*	*	*
<i>Help Others</i> (4)													*	*	%	*	*
Work Experience (5)													%	*	&	*	*
Lack of Stimulation (10)													#	&	*	*	*
Other Relatives (2)													%	*	*	*	*
Talent (7)													&	*	*	*	*
Opportunity (8)														%	*	*	*
Health (9)															%	*	*

Both the motives and influences are numbered on the left in the order they were defined in the text; motives are distinguished by italics. Each category listed on the left had more high ratings than the categories under which the symbols of levels of confidence have been placed. Thus "family" exceeded motives 1 through 6 and influences 2 through 11.

The levels of confidence of the chi squares testing the differences of the number of high, average, and low ratings are as follows: *, the 0.1 per cent level of confidence; %, the 1 per cent level; #, between the 1 and 2 per cent levels; and &, between 2 and 5 per cent levels of confidence.

4. *The Motivating Factors*

The drive for "self-expression" has not previously been mentioned as important in studies of vocational development, yet in this study it was the highest of the motives in terms of the percentage showing evidence for it, standing second, third, and sixth in the comparisons of relative importance within the three groups. Such a motive was, perhaps, a rather intangible influence, but nevertheless showed enough stability and concreteness to be described and discerned. It may well be closely related to "status," as both included striving for responsibility; probably achievement of status yields a "sense of accomplishment." It does appear that the male teachers, more than either other group, were motivated by this desire.

Another unique aspect of this study was assaying the importance of "status" to vocational orientation; usually in the past the effect of a drive for status has been inferred from children's job preferences. The results here, of course, may have reflected the bias of the raters, but nevertheless represent greater

TABLE 3
COMPARISON OF THE RELATIVE IMPORTANCE TO THE MALE TEACHERS
OF THE SEVENTEEN CATEGORIES

	Motives						Influences										
	1	2	3	4	5	6	1	2	3	4	5	6	7	8	9	10	11
Family (1)			&	*		*	*	*	%	*	*	*	*	*	*	*	*
Self-expression (5)				*		*	*	*		*	%	*	*	*	*	*	*
Status (1)				*		*	*	*		*	%	*	*	*	*	*	*
Money & Security (3)				*		*	*	*		*	*	%	*	*	*	*	*
Association people (2)				*		*	*	*		*	*	&	*	*	*	*	*
School (4)				*		*	*		%	*	*	*	*	*	*	*	*
Early Contact (11)				%		*	*			%	*	*	*	*	*	*	*
Financial Need, Lack (6)				%		%	*			*	*	*	*	*	*	*	*
Work Experience (5)				#		&	*		%	*	*	*	*	*	*	*	*
Friends (3)							*			*	*	*	*	*	*	%	*
Variety (6)							*			*	*	*	%	*	*	%	*
Help Others (4)							*			*	*	*	%	*	*	%	*
Lack of Stimulation (10)																&	*
Other Relatives (2)																	
Talent (7)																	
Opportunity (8)																	
Health (9)																	

Both the motives and influences are numbered on the left in the order they were defined in the text; motives are distinguished by italics. Each category listed on the left had more high ratings than the categories under which the symbols of levels of confidence have been placed. Thus, "family" exceeded motives 3, 4 and 6, and influences 2 through 11.

The levels of confidence of the chi squares testing the differences of the number of high, average, and low ratings are as follows: *, the 0.1 per cent level of confidence; %, the 1 per cent level; #, between the 1 and 2 per cent levels; and &, between the 2 and 5 per cent levels of confidence.

objectivity than had previously been achieved. The professional subjects clearly manifested more of this desire for status as defined here—a feeling for vocational, social, and educational standing. The finding of such contrasts might be interpreted as a contradiction of earlier studies which showed a generally excessive choice of the professional fields by grade school children. It is possible, however, that either the factory workers had forgotten those desires that they had failed to fulfill, or they may have disliked revealing them, or may have been enough restricted in verbalizing their feelings that these escaped the interviewer. Certainly very few of the subjects commented directly on the importance of prestige, although the ratings and other data showed that it was relatively high in importance.³ All in all, the drive was less crucial to the factory workers, who had to date generally achieved less status.

³Yet one subject not only commented on it, but stated that it had been his main driving force in life. He concluded, however, that once a position was attained, the prestige quickly faded.

TABLE 4
COMPARISON OF THE RELATIVE IMPORTANCE TO THE FACTORY WORKERS
OF THE SEVENTEEN CATEGORIES

	Motives						Influences										
	1	2	3	4	5	6	1	2	3	4	5	6	7	8	9	10	11
Financial need (6)	*	*	&	*	*	*	*	*	*	*	%		*	*	*	%	*
Money and Security (3)		*		*		#		*	&	*			*	*	*		*
Lack of Stimulation (10)	#	*		*		#		*		*			*	*	*		*
Work Experience (5)	%	*		*		#		*	&	*			*	*	*		*
Variety (6)		*		*				*		*			*	*	%		*
Self-Expression (5)		*		*				*		*			*	*	%		%
Family (1)		*		*				*		*			*	*	%		%
Friends (3)		%		*				%		%			*	#			&
Status (1)				*				&					%				
Association People (2)				%													
School (4)				%													
Health (9)				%													
Early Contact (11)				#													
Other Relatives (2)				&													
Opportunity (8)				&													
Talent (7)																	
Help Others (4)																	

Both the motives and influences are numbered on the left in the order they were defined in the text; motives are distinguished by italics. Each category listed on the left had more high ratings than the categories under which the symbols of levels of confidence have been placed. Thus, "financial need or lacks" exceeded all motives and all of the other 10 influences.

The levels of confidence of the chi squares testing the differences of the number of high, average, and low ratings are as follows: *, the 0.1 per cent level of confidence; %, the 1 per cent level; #, between the 1 and 2 per cent levels; and &, between the 2 and 5 per cent levels of confidence.

The male teachers showed more active interest in associating with other people than did the other groups. This was especially interesting in view of the fact that the factor was second most significant for the women teachers but only fifth for the male teachers. Although only half of the factory workers gave indication of this motive, they did not feel an aversion to people, since 98 per cent reported liking people.

The drive for money and security was somewhat more evident among the two sets of males than among the women, but even 85 per cent of the latter showed some concern about their income. As stated above, it seemed that the workers showed greater deprivation from lack of money, but the professional men evidenced more drive to seek money. On the other hand, the drive for money was, within the separate groups, relatively more important to the workers; it was second in importance to them, fourth for the male teachers, eighth for the women. Thus this drive to seek money and steady income showed considerable differential in its effect on the three groups.

TABLE 5
PERCENTAGE OF SUBJECTS RATED AS REVEALING ANY EVIDENCE AT ALL AND IMPORTANT INFLUENCE FOR EACH CATEGORY*

	Total group		Women teachers		Male teachers		Factory workers	
	Any	Important	Any	Important	Any	Important	Any	Important
<i>Motivations</i>								
Self-expression	91	38	96	25	95	61	77	19
Status	90	36	93	37	96	51	74	9
Association with People	90	31	99	28	95	52	51	0
Money and Security	90	28	85	9	92	47	95	30
Variety	74	10	72	8	68	13	86	7
Helping Others	54	10	59	7	71	17	26	2
<i>Influences</i>								
Family	96	48	100	45	100	67	84	21
Financial Need or Lack	79	34	67	19	81	40	98	51
School	77	37	88	44	87	51	42	0
Early Extended Contact	75	32	80	39	84	40	51	7
Own Work Experience	74	16	56	9	87	17	84	26
Friends	73	19	80	13	68	27	70	14
Lack of Stimulation	57	18	52	17	49	12	79	30
Talent (Consideration of)	53	4	57	5	53	5	44	0
Other Relatives	41	10	52	15	36	4	30	0
Opportunity	30	6	27	5	28	3	37	2
Health	21	2	12	0	19	3	39	5

* "Any" means ratings of one to four, "Important" means ratings of three or four.

It may be that the need for "variety" falls better as a sub-category under "self-expression." And yet on its own merits it warranted some rating for a substantial number of the subjects—over 70 per cent. Although it received about the same ratings among all three groups, its relative standing within the groups was quite varied.

The male teachers exhibited more interest in "helping others" than the workers, for whom this was the least important category. There were no questions in the interview related directly to this motive, and there might be some justification for subsuming it under the drive to associate with people. However, for other vocations it might be an important influence.

5. *External Influences*

The influence of the family was clearly the most important overall, but even here there were group differences. One new finding was the 100 per cent of both teacher groups that showed at least some influence of family. Earlier studies usually sought only the *most* important influences, so that the lower percentages given for family influence are not properly compared with this finding of 96 per cent who revealed *any* evidence of family influence. Yet these results clearly imply that whether or not the family is the most important factor to the individuals, it was the most pervasive influence, noticeable in almost every case.

One apparently unique finding was the subjects' report on their families' financial support, although this has been assumed (Ginzberg *et al.*, 2, 1950, p. 175). The percentage reporting *no* such financial support at all were 3, 15, and 74 per cent for the women teachers, male teachers and factory workers (differences significant at the 5 and 1 per cent levels of confidence). Apparently many of the women teachers would not have been teachers without family financial aid, whereas more of the male teachers got along by themselves. All in all, the data pointed to a much greater family influence than has been revealed by previous studies, and once more revealed the poor vocational orientation of the factory workers: even their families did less than those of the teachers to prod them to vocational considerations.

There was considerable difference between the degrees to which the professional and non-professional subjects showed school influence; the teachers were more frequently influenced by their teachers, courses, and the educational system. Of the teachers, the women revealed this as more important than the men. It is of note that this group of subjects, attending high school between 1924 and 1938, received (or recalled) virtually no help that could be termed formal vocational guidance. Fewer factory workers remembered especially

liking some teacher, and more of them felt limited by lack of schooling. Despite its inadequacies, then, the school system seemed to be important to the vocational development of the teachers; it could have been more effective in stimulating the factory workers. Whether or not this failure was due to a clash between these subjects and the "middle class values" usually revealed in the public schools was not revealed by the data.

The percentage of subjects who showed the effect of financial limitation may appear surprisingly high for the two groups who have completed a college education. It is possible that some of all three groups used lack of money as the excuse for giving up another field, like medicine, when in actuality they could not get the training for other reasons. On the other hand, these figures might be used to support the need for increased financial aid to education, to help more people in reaching their most-desired vocation. In the overview, lack of sufficient funds appeared as a relatively potent factor in limiting vocational development, especially for the factory workers.

The attempt to estimate the effect of "early extended contact" met with considerable success, and pointed to a type of influence which will bear further study. Refinement of the definition, and added interview questions, would doubtless help clarify the picture.

The amount of influence that the subjects' friends had on their careers was not particularly high, and yet was too great to be called negligible. It operated on some 70 per cent of the entire group, and was essentially equal in importance to each. In many ways the influence of friends appeared to be sort of a middle factor—of average importance.

For what appeared to be the first time, attention was given to whether early work experience drove the subjects *out* of a field. Thus ratings of the influence of "work experience" included both the positive and negative effects of the work. For the males, at least, this influence appeared relatively important; for the factory workers who had so little other stimulation it was relatively more so.

Some case histories were so impoverished that a category "lack of stimulation" was included which gave the rater opportunity to consider any positive statements about vocational drifting, as well as to survey the entire history for an overall picture of lack of opportunities for vocational stimulation. Even though the teachers received heavy ratings on other categories, some 50 per cent of them revealed evidence that they might have received more vocational orientation. The importance of this category to the factory workers emphasized strongly the neglect that deterred their vocational development.

It should be recalled that "talent" was the category on which the two

raters did not agree; it may be because this was the least "external" of the 11 influences. Certainly some measure of the subjects' talent would have been helpful in study of this factor, as there is no gauge of the effect a difference in intelligence has on the data obtained, nor on vocational growth.

The distinguishing feature about the category "other relatives" was that it was the only factor on which the women teachers' ratings exceeded the males'. It might well have been combined with the "family" category.

"Opportunity" was considered quite in detail, to see if "happenstances" reported were really that, or actually the result of other factors. Certainly, there must be an opening or there will not be employment. Yet the fact that the person is applying in itself indicates that some force is operative, and frequently the place where he applies is related to some previous contact. As a result of the close scrutiny (and perhaps the rater's bias), the ratings for this factor proved universally low.

"Health" has the distinction of being essentially the least important factor of the 17 considered; for only the factory workers did it exceed the ratings of any other factor. It should be recalled, of course, that good health was assumed, and only mention of limitation by health was rated. It was obvious that good health was basic for the full-time employment of these subjects.

6. *Ratings vs. Tabulations of Direct Responses*

Throughout the study it was apparent that the percentage given ratings for a category and the percentage who responded directly to questions about that category did not agree. For example, 21, 41, and 61 per cent of the three groups (women and men teachers, and factory workers) said "No" to the question "Do you feel your family influenced your jobs or interests in any way?" In contrast, none of the teachers, and only 16 per cent of the workers were given zero ratings for family influence. Likewise, 98 per cent of the workers were rated as revealing the limiting effect of financial need, but only 60 per cent said they felt limited in their job choice by anything at all. Except for the categories "health" and "other relatives," such gaps were found. These discrepancies between response to direct questions and the results of the "global judgments" emphasize the need for a method which goes beyond the subjects' report of his own understanding of influences on vocational choice.

7. *General Implications*

Although this was primarily a descriptive study that sought to fill gaps in the basic knowledge about interest development, certain implications relevant to vocational orientation appeared to be supported by the findings.

1. Apparently childhood stimulation is of vital importance to vocational orientation. The group who revealed more stimulation had reached a professional job level; those who appeared to have no stimulation appeared to have just dropped into the jobs they could find without extended planning. Some parents had apparently done their children a disservice by neglecting vocational stimulation. Such parents need reorientation themselves, or else some other agency should take over the rôle. The school had not done as much as it could, yet other community agencies are hardly in a position to accomplish as much as the school. Thus it would seem that parent education was needed, and that more vocational stimulation in the schools was called for.

2. It appeared that lack of finances was reported as preventing many subjects, both professional and non-professional, from reaching their vocational objectives. Financial assistance for education might have assisted some of that group in reaching jobs of their choice, and hence have led to greater job satisfaction.

3. It seemed quite clear that different influences and motives operated to turn the subjects into different fields. This implication is important to considerations of "class" or "level." The subjects came from relatively similar homes, yet reached quite divergent occupational levels. It appeared that those who received stimulation rose in level, whereas the others made little vocational progress. There is, of course, the possibility that the stimulation was there but some personal characteristics had blocked any effect it might have had on the one group of individuals. Or possibly the stimulation was there but the method of investigation failed to bring it out from the one group.

4. It would appear that even at the relatively advanced age of these subjects (28 to 40), there is considerable vocational reorganization in the offing. Both the upswing in interest in the late twenties and the subjects' expressed goals supported the picture of continued change in vocations.

5. Vocational choices are part of the behavior of the individual in his social group. Any understanding of the interactions involved in vocational selection contribute to an increased understanding of human behavior. Certainly in this study the rôle of the family and of early childhood experiences have been emphasized. The motivation to certain social goals was clearly demonstrated.

D. INDICATIONS FOR FURTHER STUDY

The efficacy of reviewing vocational histories obtained by interviewing different types of subjects has been shown. The major methodological changes would be in the line of expansion of the study: clear recordings of

all interviews for study by the raters; added time to probe for reasons for loss of interests; added training for interviewers and raters; inclusion of measures of intelligence and other factors. Each addition, however, should be weighed carefully in terms of the increased demands put on the subjects; each new step would limit somewhat the subject group to which the study could be applied without great increase in cost. Re-interview for check on recall accuracy should be included.

There is clearly need for further study in several areas of this work. The motives and influences found here need to be redefined and verified; their relative importance needs verification; certainly their inter-dependence should be checked to see if their number can be reduced. Furthermore, data for other occupational groups is needed, as is data for extended age levels. For example, do doctors and sales clerks show these same patterns in their development; is the importance put on finances here due to the subjects' being "depression children," or would other age groups reveal the same? For the developmental picture, much more should be done on age trends in influences.

E. SUMMARY

1. Six categories of influencing motives and 11 of external influences were drawn from the interviews, and were rated for their importance. The method of estimating the relative importance of these categories by chi square was both feasible and productive. The results of this analysis were often at variance with earlier studies: new categories were reported, and much greater percentages of the subjects revealing the influences were found. Yet the data within the study presented similar contrasts between percentages giving direct answers about, and receiving ratings on, each influence. This emphasized the need for overall estimates to overcome the individual's lack of vocational insight.

2. The results revealed many forces acting on the subjects, so that the latter cannot be expected to give a single answer as to what influenced them. The male teachers showed much the most influences to have been operative; the factory workers showed much fewer or weaker influences.

3. The ratings of the motives indicated that evidence could be gathered as to the driving forces that stimulate vocational choice. The drive for self-expression was most important of the motives that were rated. There was clear evidence to support the importance of a desire to achieve status, which was the second most important motive. Yet the factory workers revealed much less of these drives than the teachers. This was also true of the third most important motive, a need to be in contact with people. Only on the

drive to get money and security did the factory workers reveal any strong motivation.

4. The evidence was clear that the influence of family was of top importance for the teacher groups; it was found to some extent in the interview of every teacher. It was important, but not as highly so, in the factory worker ratings. The financial help given 97 per cent of the women teachers and 84 per cent of the male teachers was an obvious influence on their reaching their chosen field of work, and contrasted with the help reported by 14 per cent of the factory workers.

5. School, financial limitation, and early extended contact with some vocational area were of considerable importance, and received about the same average ratings for the 193 subjects. However, their relative importance to the three groups varied. There was indication that the school could have done more in the line of vocational orientation, especially for the factory workers.

6. The other influences received enough ratings to emphasize the fact that more than one or two factors operate in vocational choice. For individual subjects these other influences may have been most crucial. For the male subjects, work experience was relatively important. But in the total picture, the subjects' family, school, finances, and extended contact with a vocational area were the most predominant of the influences.

REFERENCES

1. BERDIE, R. F. Factors related to vocational interests. *Psychol. Bull.*, 1944, **41**, 137-157.
2. GINZBERG, E., GINSBURG, S. W., AXELRAD, S., & HERMA, J. L. The problem of occupational choice. *Amer. J. Orthopsychiat.*, 1950, **20**, 166-198.
3. NORTON, J. L. Patterns of vocational interest development and actual job choice. *J. Genet. Psychol.*, 1953, **82**, 235-262.

Graduate School
Alfred University
Alfred, New York

THE HOARDING BEHAVIOR OF THE MOUSE: I. THE RÔLE OF PREVIOUS FEEDING EXPERIENCE*

Departments of Psychology, Bucknell University and University of Maryland

WENDELL SMITH AND SHERMAN ROSS¹

A. INTRODUCTION

Recently, Beach (1) has presented a most timely paper on the status of comparative psychology, its development and decline, and its advantages and disadvantages. In criticizing current methods in animal studies, Beach decries the overemphasis on the study of a single species, and on limited topics such as learning, to the near exclusion of many important complex patterns of response.

... The growing emphasis upon learning has produced a complementary reduction in the amount of study devoted to what is generally referred to as "unlearned behavior." Any pattern of response that does not fit into the category of learned behavior as currently defined is usually classified as "unlearned" even though it has not been analyzed directly. Please note that the classification is made in strictly negative terms in spite of the fact that the positive side of the implied dichotomy is very poorly defined. Specialists in learning are not in accord as to the nature of the processes involved, nor can they agree concerning the number and kinds of learning that may occur. But in spite of which uncertainty most "learning psychologists" confidently identify a number of complex behavior patterns as "unlearned." Now the obvious question arises: Unless we know what learning is—unless we can recognize it in all of its manifestations—how in the name of common sense can we identify any reaction as "unlearned"?

The fact of the matter is that none of the responses generally classified as "instinctive" have been studied as extensively or intensively as maze learning or problem-solving behavior. Data relevant to all but a few "unlearned" reactions are too scanty to permit any definite conclusion concerning the rôle of experience in the shaping of the response. And those few cases in which an exhaustive analysis has been attempted show that the development of the behavior under scrutiny is usually more complicated than a superficial examination could possibly indicate (1, p. 121).

*Received in the Editorial Office on May 23, 1951.

¹This study was supported in part by a grant-in-aid from Bucknell University. The writers express their thanks to Mr. Norman Smith for his assistance throughout the experiment.

If we are to have a systematic comparative psychology, it seems necessary that we select these animals which promise to be most appropriate for the problem with which we are concerned. Hoarding is a behavior pattern which has generally been labeled instinctive, innate, or unlearned. This behavior needs systematic analysis since the results of such an analysis should aid in further delineating the place of experience in a so-called instinctive activity.

The laboratory rat has usually served as the subject for investigations of hoarding behavior even though the Norway rat rarely hoards food in its natural environment (2). As Beach suggests, it would be more reasonable to study hoarding in an animal which is a "natural" hoarder. Just what a "natural" hoarder is remains to be determined. For our purposes, a natural hoarder is an animal which hoards with greater relative frequency than the laboratory rat. It may also be assumed that a natural hoarder depends on this activity for survival.

A second reason for selecting a type of animal other than the rat, is to contribute further data on another species. For these reasons, and others, the mouse (Jackson Laboratory C3H) was selected for study. The mouse has been relatively little used in psychology except for the study of seizures and for the study of some aspects of social behavior such as aggression (22, 24, 25, 29). Mice have been found difficult to use in behavior studies which require discrimination and serial learning. This may be due to special unsolved motivational problems. In any event, the mouse would seem to be superior to the rat as a subject in investigating hoarding behavior in that, as we shall see, the behavior pattern can be elicited without deprivation and at a high frequency.

B. REVIEW OF THE LITERATURE

The subject of hoarding has long been of interest to biologists and students of comparative behavior. Many naturalists have described the storing habits of mammals under natural living conditions. These field studies have been of value in: (a) acquainting interested workers with the number of species which exhibit this behavior pattern, (b) describing the extent of hoarding for a given species, (c) describing the kind of materials hoarded, (d) describing the manner of hoarding in each species, (e) describing individual variations in hoarding, (f) suggesting under what conditions hoarding occurs, and (g) presenting tentative hypotheses of the factors, both situational and organismic, which might be responsible for hoarding.

Hoarding is most generally characteristic of RODENTIA, although Ingersoll (10) reports hoarding by the garden mole (order INSECTIVORA), rat kan-

garoo, wombats, and koala (order MARSUPIALIA) and the platypus (order MONOTREMATA). Many carnivores, which must range over great areas during food shortages, also store surplus foods. Foxes, bears, wolves, jackals, cats, and weasels are known to cache their surplus kill. Jays and crows have been observed to store food.

The gnawers are the true hoarders among the mammals, however. Viscacha, kangaroo rats, pocket mice, gophers, spalax, muskrat, pack rats, deer mice, chipmunks, prairie dogs, beaver, hamsters, and squirrels are all dependent, to some degree, on hoarding for survival (6, 7, 10). Many of these animals possess pouches in which they carry the materials to the "home." Those animals lacking pouches use the mouth predominantly to carry or drag the material, although the rat kangaroo uses its prehistoric tail for carrying materials (10).

The materials stored vary over a wide range. The garden mole is reputed to store worms for consumption by "biting off their heads" (6, 10). Grass, seeds, roots, thistle stalks, insects, bulbs, fruit, grain, nuts, bark, and snails are some of the subjects commonly stored by one or more members of order RODENTIA. Viscacha, gophers, pack rats, and hamsters are known to hoard materials for which there is little apparent need.

The materials hoarded are most frequently stored in or near the home. Gophers store in one central chamber reached by several tunnels. Mole rats store in closets located along runways under the sand. Prairie dogs utilize multiple chambers for storing fodder and refuse. Viscacha drag the objects to the hillock around the entrance to the burrow. Beavers fasten limbs to the bottom of the stream near the entrance to the lodge or tunnel and cut off sections to carry inside when the stream is covered with ice. Squirrels bury food in and around the nest. The carnivores usually drag the surplus to a safe place and either bury or cover it. In every case the surplus food is deposited in what appears to be the place of maximum security from enemies for the animal. The factor of availability, as well, must be considered.

Quantitative investigations of hoarding behavior are necessary if we are to add to the fund of knowledge given us by the naturalists. Several common laboratory animals are available for research in this area, e.g., rats, mice, and hamsters. In 1939, Wolfe (30) reported the first quantitative investigation of the hoarding of food pellets by the laboratory rat. Many other papers have appeared in the literature on this subject since that time. Two investigations have dealt with the hamster (20, 21), and one incidental observation has been reported on the mouse by Scott (23). Scott reported

that he had observed no hoarding by mice even when the food pellets were presented in a small size in a food hopper in the center of a multiple escape pen. The mice nested in boxes which were located approximately 20 inches from the food hopper, each box being connected to the center of the pen by an alley. From Scott's description of the situation, it would appear that the conditions for eliciting food hoarding were present.

C. HYPOTHESES USED IN THE EXPLANATION OF HOARDING

Several possible explanations for hoarding have been suggested. These hypotheses may conveniently be classified as those which attribute hoarding to (a) environmental adaptation (natural selection), (b) as instinct, (c) a physiological deficit of some kind, and (d) learning.

Ingersoll (10) and Hamilton (6) tend to consider hoarding behavior from the point of view of its survival value to the organism. Those animals which are unable to migrate and which do not hibernate during extremely cold or hot seasons must adapt in some other manner. Those which store food "accidentally" during a period of plenty will survive, and those which fail to store food will die. Over a period of several generations those characteristics which led to the survival of the parents will be transmitted to the offspring and a food gathering habit will be resultant according to this view. Ingersoll (10) writes that:

... The search for food is the foremost anxiety and occupation of these little creatures, it (the anxiety) would be increasingly stimulated as the ripening season of the seeds and nuts on which they depend advanced, and the impulse to incessant industry, so necessary in the poorer parts of the year, would not be overworked, and each animal in his haste to be up and doing, would constantly bring home more food than would be consumed, so that it would pile up in the accustomed "dining room." The gradual failure of outdoor supplies, as winter came on, would lead to the eating, with increasing frequency, of those fragments casually saved in and about the burrow or which, from their nature, would not have decayed. The animal which had been most busy and clever in food gathering would own the largest amount of the leavings of these autumnal feasts. Having the most food he would be among those of the colony or neighborhood strongest and most likely to survive, and give to his offspring the tendency to strength and industry which had been his salvation. This would be continued and shaped by the process of natural selection into a valuable, instinctive habit of gathering food in large quantities as winter provender (10, p. 452).

This same writer attributes the development of cheek pouches in some animals to the need to carry large quantities of food for eating and storage!

The theory presented by McDougall (13) is representative of those which fall into the instinct classification. The fundamental notions of this hypothesis may be seen from the following quotation:

... Some of the mammals gather surplus food and hoard it instinctively. ... Among those mammals which have fixed abodes and which feed on grain, nuts, and other such durable vegetable products, the tendency is wide-spread.

It may be that it is an instinct in process of differentiation from the feeding instinct, and that various species illustrate stages of this differentiation ... in general animals that have any kind of lair or home are apt to resent the intrusion of other creatures; and this defensive attitude is very commonly extended to a certain area surrounding "the home." ... This possessive behavior seems to be instinctive. It is certainly closely allied with hoarding. ... I am inclined to suppose that there are too distinct instincts, and that both are constituents of normal human nature. But it is so difficult to distinguish their influence in human behavior and experience that for the present we may with advantage content ourselves with the recognition of an *acquisitive* instinct; and we may attribute to it both hoarding and the defensive possessive behavior (13, p. 161-162).

With some change in terminology. MacDougall's thesis has a distinctly modern flavor. Although his concept of instinct is not acceptable to most American psychologists, it is interesting to note that Morgan (16), in a review of the literature on hoarding, entitles his paper, "The Hoarding Instinct."

A third hypothesis relates hoarding to some kind of physiological deficit brought about by previous nutritional deprivation. The work with laboratory rats by Morgan *et al.* (17) and Smith and Ross (19) has suggested that sometimes deprivation is a necessary condition for the continuation of hoarding. Morgan and his co-workers consider hoarding as primarily "the result of food deprivation," and attribute it to "... the creation of some deficit in the body which requires considerably more food deprivation than does hunger reaching a crucial level, and takes a greater amount of food intake to relieve than does hunger (17, p. 293)."

Wolfe (30) emphasized the importance of inadequate diet and related tissues needs to hoarding.

... The genesis of food storing is closely related to diet, but once started it persists through to a lesser extent even after the diet has been adequate for as long as twenty days.

The fact that rats raised on a diet sufficient in quality to enable them to reach maturity, reproduce, and appear normal in every way, will

readily store food when non-hungry white rats given a more complete diet will not, indicates that the absence of certain food elements provides distinct motivation for specific behavior (10, p. 107).

Stellar (26, 27) has attempted to study the nature of the deficit involved in hoarding and presents this hypothesis:

... Changes in the liver which occur as the result of infantile feeding frustration persist until adult life and make for defective regulation of carbohydrates metabolism in the face of hunger. Then it is assumed that the internal stimuli arising from a low level of carbohydrate balance are the determinants of the hoarding behavior (26, p. 85).

Injection of insulin and glucose failed to change the hoarding level. Stellar expressed some doubt concerning the adequacy of the experimental procedure, hence did not report any conclusions concerning the physiological mechanisms involved.

Bindra (2, 3, 4, 5), has taken a position which lies somewhere between that of the *deficit hypothesis* group (Stellar, Morgan, Johnson, Wolfe) and the *learning hypothesis* group represented by Hunt (8, 9) and Marx (11, 12). Bindra (4) presents as his hypothesis the following statement:

... If a rat is in such a state that, in its home cage, it would eat a particular food, then the rat is in such a state that in a hoarding situation, it will hoard that food. The more food that the rat would eat if in its home cage, the more food it will hoard when in the hoarding situation (4, p. 215).

Bindra (2) found that simultaneous deprivation of both protein and carbohydrate is related to hoarding. This finding clearly supports a deficit hypothesis. Deprivation of fat, vitamins, or mineral salts did not lead to hoarding. Apparently the need for only certain specific food elements will arouse hoarding.

Bindra (5) has also proposed a *security hypothesis* under which a variable of *shyness* accounts for the appearance of hoarding. This is based on a formulation by Miller and Viek (14):

... The rat may hoard to the cage because it regards the cage as a point of maximum security in the hoarding situation. That is, if the home and cage and pellets are regarded as having positive valence for the rat, and the unfamiliar alley as having negative valence, hoarding behavior may simply be the rat's solution to a geographical problem. By combining the positive valences the rat is able to achieve maximum security and minimum anxiety in the situation (14, p. 229).

McDougall (13), Ingersoll (10), and Hamilton (6) have also observed

the importance of bringing the food to a safe place for consumption. The prairie dog apparently moves his "home" frequently to avoid the necessity of foraging over great distances. Possibly there is a relationship between the hoarding of various objects and the food supply in the animal's territory. When the food supply necessitates extensive travel, the "home" is moved and hoarding to the new home starts again. Under natural living conditions, many mammals tend to eat in a relatively secure situation. This obviously necessitates carrying or transporting the food in some manner.

Hunt (8) has presented a hypothesis of hoarding based on learning theory. Hunt's explanation is particularly applicable to adult hoarding by animals which had experienced varying degrees of deprivation during infancy. Hunt

... assumed that the hunger of the infantile feeding-frustration, to which the experimental animals were subjected, aroused autonomic and humoral responses which served to energize the animals. During the period of the feeding-frustration, it was noted that the experimental animals were much more active than the controls. He further assumed that the cue-stimuli of any mild hunger could serve as conditioned stimuli to rearouse these energizing responses. On the basis of this assumption, the adult feeding-frustration was seen to differentiate between experimental and control animals because the cues involved in this relatively mild hunger rearoused the extra energizing responses in the experimental animals while it set up only a primary hunger drive in the control animals. Thus, the experimental animals would have entered the hoarding tests energized by both the conditioned energizing responses and a primary hunger drive, while the control animals entered them with only the primary hunger drive (9, p. 292).

This theory makes hoarding an acquired drive, and as has been pointed out elsewhere (9), is similar to the principle of associative reinforcement.

Marx (11, 12) has presented an analysis of hoarding based on reinforcement learning theory. This analysis is a comprehensive and fruitful treatment and deserves careful and thoughtful attention from investigators who are interested in further study of this complex, heterogeneous activity.

Marx's hypothesis is, "that the typical hoarding habit may be primarily explained as the result of the fusion of independently acquired habits" (12, p. 81). This hypothesis is based on a set of observations of the behavior of the rat in its laboratory environment. The independently acquired habits to which Marx refers are purported to be: (a) a pellet seizure response, (b) a pellet carrying and running response, (c) a pellet release response, and, (d) a "homing" habit. These four responses are said to be learned by the rat prior to the actual onset of hoarding. The responses are typically

acquired in the following manner: For rats fed a pellet diet following weaning, the response of seizing the pellet is readily acquired and constantly reinforced by the "*response produced* stimulation that invariably follows jaw muscle contraction, including those which are actually involved in direct feeding" (12, p. 82). Pellet carrying is acquired early in the animal's life as a result of "anxiety-producing" experiences such as (a) competition for the pellets from other animals, and (b) disturbance of the "home" situation by laboratory noises, attendant's hands in the cages, and the like. The habit of running with the pellet is said to be "conditioned to kinesthetic stimulation produced by pellet seizure" (12, p. 83). The running to a "safe" place in the home serves as a mechanism for reducing "anxiety." Once a "safe" place is reached, the pellet must be released if the animal is to feed. Marx says the "act of releasing the pellet may thus be considered to be conditioned to the proprioceptive stimulation associated with the cessation of movement and to be reinforced by the more or less immediately occurring direct feeding with resultant drive reduction" (12, p. 83). The final habit acquired before the onset of hoarding is learning to leave and return to the cage. This may be "explained upon the basis of the secondary reinforcing properties of the various stimuli produced by the cage (mainly olfactory and visual) and previously associated with such primary need reduction as has been effected by regular eating and drinking in the cage" (12, p. 83).

The integration of these four habits is explained by assuming that the food pellet plays a dual stimulating and reinforcing rôle. The rat enters the alley, eventually travels to the food bin, finds pellets there, and under these new anxiety conditions repeats, in order: pellet seizure, pellet carrying, pellet release, and feeding. All of these responses were previously acquired, according to this analysis, and are reactivated by the anxiety conditions associated with the relative "insecurity" of the hoarding alley. In the typical hoarding experiment, food deprivation is associated with this complex experience, and this factor increases activity while it also undoubtedly increases "anxiety."

Following the leads provided by his analysis, Marx proceeded to demonstrate other characteristics of the learning process such as generalization, extinction, and spontaneous recovery.

It seems pertinent to ask why previous investigators have failed to hit upon a similar explanation. Marx's answer is that they did not look for it, or, at least, did not look for it in the right place. A critical examination of Marx's analysis seems to be in order.

The key element in the hoarding process as described by Marx, appears to be the pre-hoarding pellet experiences, i.e., pellet seizures, pellet carrying, and pellet release. If all of these experiences were eliminated, and if hoarding is learned in the manner described by Marx, it should follow that hoarding would appear slowly, if at all, in animals lacking (deprived of) these primary experiences. The only previous investigation bearing on this point was conducted by Wolfe (30) in whose study some rats were raised on powdered food from weaning. These rats were retarded in hoarding frequency on the first day. Marx in considering this finding has the following to say:

... The appreciable increment in hoarding shown on the subsequent test days may be explained on the basis of a relatively rapid learning of responses to the food pellet. Approach and manipulatory responses would be expected on even the first encounter of such rats with a food pellet. They would result from the generalization to be expected on the basis of the olfactory and visual (texture) similarity between the food pellets and the powdered food, and would account for the slight amount of first hoarding which Wolfe found (13, p. 86).

Possibly, there should be some doubt concerning the rapid learning of a pattern of behavior as complex as hoarding, involving as it does the integration of a series of responses. In any event, it seemed worthwhile to investigate the hypothesis that *mice which are deprived of pellet experiences prior to the hoarding trials will be inferior in hoarding to mice which have had prior pellet experience.*

In view of the fact that this was an initial study of mouse hoarding, it seemed advantageous to explore further the nature of this behavior.

D. PLAN OF THE EXPERIMENT AND PROCEDURE

The plan of the experiment is relatively simple. The control group of 30 mice and the experimental group of 30 mice followed the identical procedure as follows:

Period I.—10 days of deprivation hoarding.

Period II.—15 days of satiation hoarding.

Period III.—10 days of deprivation hoarding.

The purpose of the experiment was to compare a group of mice denied previous pellet experience with a group raised in the standard manner, that is, fed laboratory pellets to which they had easy access for seizing, carrying, and releasing.

As this appeared to be an initial investigation of the hoarding behavior

of mice, the experiment was designed also to study the relative effects of food deprivation and satiation on hoarding by mice. With this design, it is possible to compare the behavior of mice in this activity, with that exhibited by rats and hamsters under laboratory conditions.

1. *Equipment*

Batteries of cages containing nine individual living units were constructed. Each living unit was 4 x 5 x 4 inches and contained an exit large enough for the mouse to enter and leave the cage with ease. The top of each individual unit was hinged for ease in feeding and removing the mouse, and as an aid in removing access pellets.

The hoarding alleys were constructed of a plywood bottom, with No. 4 hardware cloth sides and top. The alleys were 3 x 12 x 3 inches, permitting the mouse to move about with relative ease, yet small enough to discourage excessive exploration.

2. *Subjects*

The subjects for the experiment were 60 C3H mice, approximately five months old, from the colony of the Department of Psychology at Bucknell University. There were 30 males and 30 females in the group. Following weaning, the subjects were divided into two groups; an experimental group containing 15 males and 15 females, and a control group containing the same number of males and females. The males were caged in two groups of 15 each. The controls were put on a pellet diet and cared for in the standard manner. The experimental group never experienced pellet feeding since all pellets were removed from the mothers' cages following birth and a liquid diet, to be described below, was substituted. The first contact that the experimental subjects had with food pellets occurred in the hoarding alleys on the first hoarding trial.

3. *Diet*

A liquid diet of Purina fox chow pellets reduced to a meal and mixed with water, was supplied by bottle and metal tube to all experimental subjects. The bottles were green glass which served to change the visual appearance of the food, making it extremely dark in color to the experimenters. A drop of food was always present at the lip of the feeding tube, hence visual experience of the food was possible at this point. Olfactory cues were probably similar to those from the solid pellet. Frequent weighing of the animals fed this diet indicated weight increase equal to the controls.

4. Procedure

Four days before the start of the hoarding trials, the animals were transferred to the individual living cages. On the basis of prior experience with rats, hamsters, and mice, it was felt that this was a sufficient period for the subjects to become accustomed to isolation in the relatively small cage. The mice were handled at all times on a sheet of hardware cloth to minimize the excitement which occurs if mice are lifted by the tails with an accompanying loss of support.

Pellets of Purina Laboratory checkers (fox chow) were supplied to the controls *ad libitum* during the first two days of the adaptation period while the test diet was supplied to the experimentals via bottle and feeding tube.

Forty-eight hours prior to the first hoarding trial, deprivation conditions were established. Deprivation consisted of limiting all subjects to sufficient food for 12 hours, i.e., each subject was tested after approximately 12 hours of food deprivation. Previous experience contra-indicated a 23-hour hunger drive as being too severe for mice.

On the first day of the 10-day deprivation period, *Period I*, the hoarding alleys were placed before the exits of the cages, and the mice were given access to the alley and food bin for 20 minutes. A 20-minute period was used as preliminary trials with other mice indicated this to be the maximum time spent in hoarding by the majority of the group. At the end of the trial, all pellets were removed from the cages and counted. Each animal was then given sufficient food for 12 hours. Water was supplied *ad libitum*.

Following the tenth deprivation trial of *Period I*, a full ration was supplied to all subjects in preparation for *Period II*, 15 days of satiation hoarding. Full feeding was maintained until the 15th day of *Period II*, when deprivation conditions were re-instituted for 10 days, as in *Period I*. All trials were made under as nearly identical conditions as possible. The temperature of the room in which the animals were housed ranged from 72°F to 80°F. Lighting was supplied by a bank of eight, 20-watt, shaded, fluorescent bulbs suspended approximately four feet above the cages. All trials were conducted on successive mornings at nine o'clock.

E. RESULTS AND DISCUSSION

One comparison of the results for the experimental and control groups is given in Figure 1, which shows the total number of pellets hoarded for each day of the experiment by each group. If there is any significant difference between those mice fed a solid pellet diet and those fed a liquid diet, it should be most apparent in the early trials. An inspection of the figure

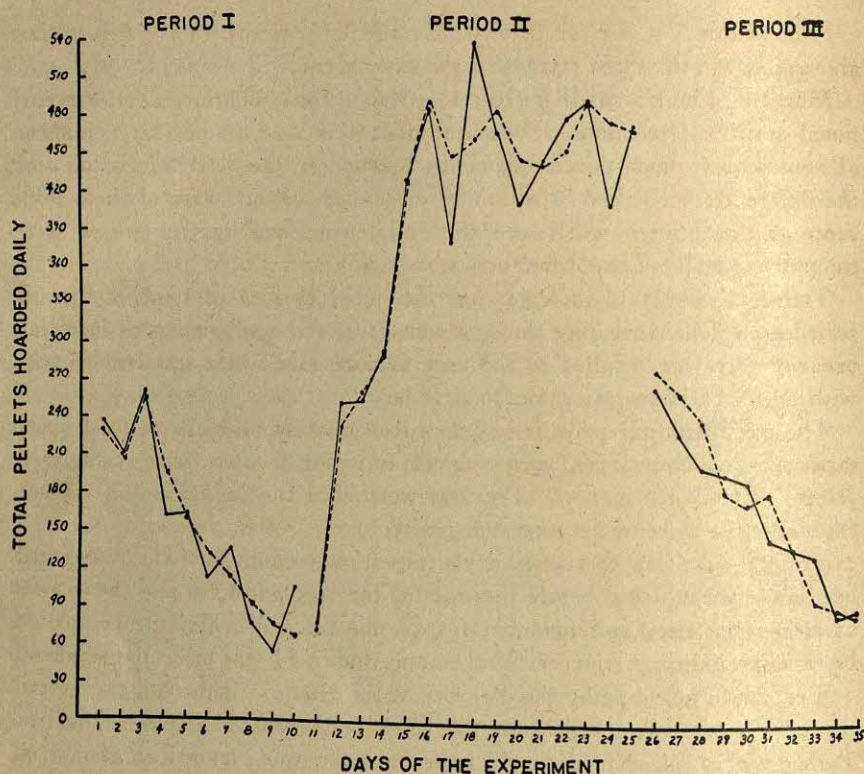


FIGURE 1

TOTAL NUMBER OF PELLETS HOARDED FOR EACH DAY OF THE EXPERIMENT BY THE CONTROL GROUP (SOLID LINE) AND THE EXPERIMENTAL GROUP (DASHED LINE)

reveals that curves for the two groups are highly similar, not only in the initial trials, but throughout the 35 days of hoarding tests. The experimental group hoarded a total of 228 pellets on the first day while the control group hoarded 237 pellets. The difference of nine pellets is not statistically significant.

TABLE 1

SHOWING THE AVERAGE DAILY NUMBER OF PELLETS HOARDED BY MALES, FEMALES AND ALL SUBJECTS IN THE CONTROL AND EXPERIMENTAL GROUPS FOR THE THREE PERIODS OF THE EXPERIMENT

	Period I (Deprivation)		Period II (Satiation)		Period III (Deprivation)	
	Exper.	Control	Exper.	Control	Exper.	Control
Males	5.94	5.82	11.13	10.98	4.62	4.30
Females	4.06	4.47	15.39	16.47	6.94	7.43
All	5.07	5.21	13.12	13.44	5.70	5.72

From Table 1 it is clear that the average daily means for the two groups are similar for the three periods of the experiment.

Morgan (15) has made it clear that it is not legitimate to apply conventional statistical techniques to hoarding data because of the large number of zero scores which occur. For this reason, the levels of significance for the differences in Table 1 are not reported. However, *t*-test of the significance of the differences between the experimental and control groups were computed. None of the differences were significant.

The data appear to indicate that the mice deprived of some of the experiences which Marx reports as important in the development of hoarding behavior were not inferior in hoarding to mice raised in a standard manner with adequate experience with food pellets.

The control group subjects had approximately four months of food pellet experience which provided ample opportunity for them to seize, carry, and release pellets each day. The pellets were placed in an open dish in each cage and the mice were observed removing the pellets from the dish and carrying them to another part of the cage before eating. If the pellets had been provided in a feeder which prevented carrying and releasing, the absence of clear-cut differences between the experimental and control groups would be easier to interpret, however, this was not the case in the present experiment.

The factor of learning has not been eliminated by depriving the experimental group of the opportunity for seizing, carrying, and releasing pellets. Perhaps it is possible that the experimental group learned to hoard as frequently as the controls on the first trial. As Marx (12) has pointed, olfactory, gustatory, and visual experiences in the home cage could serve as cues from which the mice might generalize as to the hoarding situation. Visual cues in the home cage were reduced for our experimental group but there were ample olfactory and gustatory experiences.

The alleys used in this study were relatively short (12 inches long) and narrow (3 inches wide), hence, exploration was discouraged. Short alleys may serve to mask the rapid learning of the hoarding pattern, according to Marx's analysis. If we had made detailed comparative observations of the initial approaches of the experimental and control groups to the hoarding situation, we would be in a better position to gauge the importance of alley length on time of the pellet return following the first alley entrance. Such observations were made on only five mice in each group.

The length of time for the first full entrance into the alley was estimated for each of these sub-groups. The number of partial entrances was recorded also. The results indicated a slightly longer latency of alley entrance, and

more partial entrances by the control subjects than by the experimentals. However, these observations were made on small numbers, and in a somewhat subjective manner, hence, it would be unwise to consider them more than suggestive. Perhaps it is possible that less of what Marx calls "anxiety" was produced in the experimental group by bottle feeding than was produced in the control group by pellet feeding. The dropping of pellets into the cages of the controls and the resulting competition in securing and eating the pellets might serve to produce anxiety. In any event, it appears that the experimental subjects approached the hoarding in a manner not unlike that of the controls. Longer alleys and the equation of "anxiety-producing" situations might lead to different results.

The results of the present experiment may be used to support some hypothesis of hoarding other than that of reinforcement learning theory. The majority of writers on this subject have tended to support some type of deficit hypothesis, although they have not presented any detailed analysis of the manner in which a deficit leads to the particular series of activities which has been called hoarding behavior. Marx seems to be the first investigator to concern himself with analysis of the overt behavior. Marx's analysis of the behavioral sequence may be found to be faulty, but he has succeeded in calling attention to the need for an experimental analysis of the total complex, heterogeneous activity.

F. GENERAL CHARACTERISTICS OF MOUSE HOARDING

Figure 1 indicates some other factors in hoarding by mice which are of more general interest. First, it may be seen that there is a decided trend toward the hoarding of a decreasing number of pellets during Period I. It is apparent also that fewer and fewer animals attained a criterion of four pellets hoarded per trial until by the ninth day of the deprivation trials only 18 per cent of the males and 30 per cent of the females were still hoarding.

This result seems to follow logically, i.e., we would expect that as the animals became hungrier, they would eat more and hoard less. However, in general, the results from hoarding investigations with rats do not support this logical conclusion. Even when rats are fed after the hoarding trials as were these mice, the frequency of hoarding tends to increase gradually, reach a peak, and level off (18, 19). What is exhibited in the present experiment is the decreasing frequency of hoarding which has been called hoarding extinction in rats and hamsters. This trend for rats and hamsters, however, usually occurs during satiation, not during deprivation. Other experiments confirm this difference for mice.

It may be seen from Figure 1 that during the 15 days of satiation trials, hoarding scores increased throughout the period, and no extinction of the response is evident. At the peak of the curve, 92 per cent of the females and 87 per cent of the males were attaining or surpassing the criterion of four pellets hoarded.

The second 10 days of deprivation (*Period III*) gave about the same results as *Period I*. In this case, at the low point, 53 per cent of the females (as compared to 30 per cent in *Period I*) and 25 per cent of the males (as compared to 17 per cent in *Period I*) attained or surpassed the criterion.

These results seem to support the conclusion that mice hoard with greater frequency when satiated than when deprived. We may ask whether mice which had never been deprived (insofar as could be determined) would hoard at a relatively high frequency. This question will be dealt with in another paper.

1. Sex Differences

Whether or not there are reliable sex differences in hoarding has not been clearly answered in previous investigations. Reference to Table 1 indicates that the females hoarded more than the males during *Periods II* and *III*. More males than females failed to meet criterion during *Periods II* and *III*. It is possible that females are more affected by deprivation, and "rebound" more intensely during satiation. Marx (12) has suggested that a greater frequency of hoarding by females may be related to the estrus cycle. Presumably, the hoarding habit becomes conditioned to, and is then activated by, the sex drive.

2. Qualitative Observations

Previous writers on the subject of hoarding have usually described the hoarding behavior as a stereotyped pattern of responses. The animal runs out of the cage, down the alley, picks up a pellet in its mouth, turns around and runs back to the cage where it deposits the pellet and then starts the sequence over again. In general, this is a fairly good description of the rat while it is hoarding.

Mice exhibit a great deal more variability both in the day-to-day responses and in the differences between individuals, hence, it seems appropriate to describe in some detail the variations in this behavior.

For the first two or three days the animals are extremely active and excited. Tail rattle and "hip wiggle" occur frequently during the first half of the hoarding period. This behavior is indicative of excitement and aggression. As no "dry runs" were experienced by the mice, this initial excitement is probably part of the process of adaptation to a novel situation.

Within the first 10 minutes of the initial trial, the majority of the mice begin to hoard. The most frequent sequence of response is: (a) climbing up on the pile of pellets, (b) digging into the pile, (c) pushing back the pellets with the forepaws, and then forcefully with the hind legs, and (d) finally selecting a pellet and carrying it in the mouth into the cage. This sequence may go on regularly for as long as 10 minutes by which time the pile of pellets is spread over much of the floor space of the alley, and the end of the alley where the pellets were piled away may be quite bare. To the observer, some mice appear to be "seeking" something. Mice that did not hoard to the criterion level at all, attempted to clear away the pile of pellets from the end of the alley and when this was accomplished would then engage in some other activity. Once the mice had hoarded to "satiation" for the period, the remainder of the time was spent in one or more of the following activities: (a) climbing the sides of the alley, (b) biting the wire mesh, (c) attempting to escape at the junction of the alley and the cage, (d) eating, (e) carrying a few more pellets into the cage at the end of the period, (f) grooming, and (g) occasionally sleeping.

Some mice do not disturb the pile markedly, but carry from the bottom of the pile as do rats and hamsters ordinarily. Some of the mice adopt novel techniques of transporting the pellets. They kick them with their hind legs or push them with forepaws and noses until they reach the threshold of the cage entrance. Then, they may lift the pellet in from the outside of the cage, or stand in the cage entrance with the body inside and lift the pellets in. A few mice have been observed to pick the pellets up at the cage entrance and toss it over the shoulder. Still another procedure is for the mouse to carry the pellet from pile to entrance, and then from entrance into the cage. A given animal may use all these techniques of transporting (carrying, kicking, pushing, and lifting) the pellets in a single hoarding period.

The mice that use the pushing and/or kicking methods most frequently, are generally inefficient hoarders. They seem to work harder but accomplish less than the carriers and lifters. Pushing and kicking are seen frequently in the activities of the "non-hoarders," i.e., those mice which fail to meet the criterion of four pellets hoarded per trial more than once or twice in 15-20 trials. It is of some interest to note that during satiation trials the "non-hoarders" are usually as active as the hoarders, but do not carry pellets into the living cages. Two of the "non-hoarders" in this experiment piled pellets near the cage entrance almost every day but rarely carried more than one or two inside. At other times, the non-hoarders dig into the

pile and move the pellets around but do not complete the pattern. It is highly probable that the majority of "non-hoarders" are, in reality, inefficient hoarders. During deprivation tests, "non-hoarders" tended to eat for a slightly longer period of time than did the hoarders.

Some of the mice become very disturbed when the excess pellets were removed from their cage following the hoarding trial. They ran around the cage rapidly. Tail rattles were frequent. Defensive postures occurred, and there were aggressive acts such as biting and striking at the tweezers used to remove the pellets. In the course of the excitement, a few mice picked up pellets in their mouths and held them up high. If the experimenter removed the pellet from the mouth, the mouse frequently would pick up another pellet and hold it in the same position. Sometimes, a mouse would push and/or carry the pellets to the end of the cage where the pellets were being removed.

There does not appear to be any established manner of storing the pellets in the cages. They are usually piled toward the back and along the sides of the cage leaving the center and entrance clear. Occasionally, all pellets are piled near the front of the cage by a given subject, and the mouse then climbs over the pile to gain access to the alley.

There does not appear to be any established manner of storing the pellets in the cages. They are usually piled toward the back and along the sides of the cage leaving the center and entrance clear. Occasionally, all pellets are piled near the front of the cage by a given subject, and the mouse then

In summary, hoarding by mice is associated with great activity, particularly during the first half of the hoarding period. A wide variability in the manner in which the mouse hoards is evident. Hoarding is, in part, an individualized performance, and cannot be considered invariant. We may well ask, as Bindra (3) did, "how does such a wide range of physiological and environmental conditions lead to the same pattern of behavior (p. 155)," and how do so many patterns of behavior occur in such a standardized situation?

G. SUMMARY

The experiment reported deals with the effect of previous feeding experience on hoarding in the mouse (C3H, JAX).

A total of 60 subjects (30 males 30 females) were divided into two equal groups by number and sex. The control group was fed fox-chow pellets from weaning and hoarded these pellets during the period of the experiment (35 days). The experimental group was fed a prepared liquid mash of the pellets with water, but had no experience with food pellets from

weaning. The experiment proper was started when the mice were approximately five months old.

The pellet hoarding tests were conducted for 35 consecutive days for both groups under conditions of deprivation (10 days), satiation (15 days) and deprivation (10 days). The results are compared for each of the groups. No significant difference appears in the frequency of hoarding for either of the groups. These results are considered in terms of various hypotheses of hoarding behavior, giving particular attention to the stimulus-response reinforcement approach.

A general description is also presented of mouse hoarding, including hoarding-deprivation relationships, sex differences, and qualitative observations.

REFERENCES

1. BEACH, F. A. The Snark was a Boojum. *Amer. Psychol.*, 1950, **5**, 115-124.
2. BINDRA, D. Hoarding behavior of rats: Nutritional and psychological factors. Doctoral dissertation, Harvard University Library, 1947.
3. ———. Water-hoarding in rats. *J. Comp. & Physiol. Psychol.*, 1947, **40**, 149-155.
4. ———. The nature of motivation for hoarding food. *J. Comp. & Physiol. Psychol.*, 1948, **41**, 211-218.
5. ———. What makes rats hoard? *J. Comp. & Psychol.*, 1948, **41**, 397-402.
6. HAMILTON, W. J., JR. American Mammals: Their Lives, Habits, and Economic Relations. New York: McGraw-Hill, 1939. (Chap. VI.)
7. ———. The mammals of Eastern United States. (Vol. II.) Ithaca, New York: Comstock, 1943.
8. HUNT, J. McV. The effects of infant feeding-frustration upon adult hoarding in the albino rat. *J. Abn. & Soc. Psychol.*, 1941, **36**, 338-360.
9. HUNT, J. McV, SCHLOSBERG, H., SOLOMON, R. L., & STELLAR, E. Studies of the effects of infantile experience on adult behavior in rats: I. Effect of infantile feeding frustration on adult hoarding. *J. Comp. & Physiol. Psychol.*, 1947, **40**, 291-304.
10. INGERSOLL, E. The Life of Mammals. New York: Macmillan, 1917.
11. MARX, M. H. Experimental analysis of the hoarding habit in the rat: I. Preliminary observations. *J. Comp. & Physiol. Psychol.*, 1950, **43**, 295-308.
12. ———. A stimulus-response analysis of the hoarding habit in the rat. *Psychol. Rev.*, 1950, **57**, 80-94.
13. McDUGALL, W. An Introduction to Social Psychology. Boston: Luce, 1923.
14. MILLER, G. A., & VIEK, P. An analysis of the rat's responses to unfamiliar aspects of the hoarding situation. *J. Comp. Psychol.*, 1944, **37**, 221-231.
15. MORGAN, C. T. The statistical treatment of hoarding data. *J. Comp. Psychol.*, 1945, **38**, 247-256.
16. ———. The hoarding instinct. *Psychol. Rev.* 1947, **54**, 335-341.
17. MORGAN, C. T., STELLAR, E., & JOHNSON, O. Food deprivation and hoarding in rats. *J. Comp. Psychol.*, 1943, **35**, 275-295.
18. ROSS, S., SMITH, W. I., & NIENSTEDT, C. W., JR. The hoarding of non-relevant material by the white rat. *J. Comp. & Physiol. Psychol.*, 1950, **43**, 217-225.

19. SMITH, W. I., & ROSS, S. The effect of continual food deprivation on hoarding in the albino rat. *J. Genet. Psychol.*, 1950, **77**, 117-121.
20. ———. Hoarding behavior in the golden hamster (*mesocricetus auratus*.) *J. Genet. Psychol.*, 1950, **77**, 211-215.
21. ———. A note on food hoarding by adrenalectomized hamsters. (In preparation.)
22. SCOTT, J. P. Genetic differences in the social behavior of inbred strains of mice. *J. Hered.*, 1942, **33**, 11-15.
23. ———. Social behavior, range, and territoriality in domestic mice. *Proc. Indiana Acad. Sci.*, 1944, **53**, 188-195.
24. ———. Group formation determined by social behavior; a comparative study of two mammalian societies. *Sociometry*, 1945, **8**, 42-52.
25. ———. "Emotional" behavior of fighting mice caused by conflict between weak stimulatory and weak inhibitory training. *J. Comp. & Physiol. Psychol.*, 1947, **40**, 275-282.
26. STELLAR, E. Studies of the effects of infantile feeding frustration on the adult hoarding of the white rat. Master's thesis, Brown University Library, 1942.
27. ———. The effect of epinephrine, insulin, and glucose upon hoarding rats. *J. Comp. Psychol.*, 1943, **36**, 21-31.
28. STELLAR, E., & MORGAN, C. T. The rôles of experience and deprivation in the onset of hoarding behavior in the rat. *J. Comp. Psychol.*, 1943, **36**, 47-55.
29. UHRICH, J. The social hierarchy in albino mice. *J. Comp. Psychol.*, 1938, **25**, 373-413.
30. WOLFE, J. B. An exploratory study in food storing in rats. *J. Comp. Psychol.*, 1939, **28**, 97-101.

Department of Psychology
Bucknell University
Lewisburg, Pennsylvania

Department of Psychology
University of Maryland
College Park, Maryland

THE HOARDING BEHAVIOR OF THE MOUSE: II. THE RÔLE OF DEPRIVATION, SATIATION, AND STRESS*

Departments of Psychology, University of Maryland and Bucknell University

SHERMAN ROSS AND WENDELL SMITH¹

A. INTRODUCTION

In a previous study of hoarding by C3H mice (12), it was found that mild food deprivation led to a decreased frequency of hoarding, while satiation was accompanied by a relatively high level of hoarding. Since this result is contrary to the findings for rats (7,9) and hamsters (10, 11), further investigation of the rôles of deprivation and satiation seemed advisable.

A study of the frequency of the hoarding of food pellets by non-deprived animals would seem to be of value in delineating the rôle of deprivation. Reviews of the literature on the hoarding behavior of rats by Morgan (6) and the Committee of Food Habits of the National Research Council (2) have assigned a major rôle to deprivation. In the N.R.C. review, it is stated that deprivation, "... appears to be the most important factor contributing to hoarding behavior ... (2, p. 73)."

Some hoarding by non-deprived rats has been reported, but it has been generally of lower frequency than that exhibited by rats subjected to food deprivation. Bindra (1) found that non-deprived rats would hoard sweetened-mash. Licklider and Licklider (3) report hoarding of plain and foil covered food pellets by rats which were weaned on food pellets and not subjected to deprivation. In the Lickliders' study deprivation hoarding trials were introduced on the 40th day of the experiment. The rats hoarded more on the test trials than they had ever hoarded under satiation conditions. Licklider proposes that the perceptual factors in the hoarding situation are the important determinants in the process. The rôle of deprivation lies in the new influence which it exerts on the perceptual processes (3, p. 133). This is an interesting approach to the study of hoarding since it changes the emphasis from the search for a deficit which releases a neural mechanism to the overt behavioral factors involved in this complex process.

*Received in the Editorial Office on May 23, 1951.

¹This study was supported in part by a grant-in-aid from Bucknell University. The writers wish to express their thanks to Mr. Norman Smith for his assistance in making and recording observations.

If hoarding frequency can be shown to be higher in mice which have experienced no known food deprivation than in those which have experienced deprivation, the rôles of deprivation in the hoarding process would appear to be inhibiting rather than facilitating.

A factor which seems to have been overlooked in comparing deprived and satiated animals is the basic adequacy of the diet. Does the pellet food which is used as the stock diet for many rodent colonies satisfy the nutritional needs of the animals? Are the "satiated" animals actually *satiated*? A comparison of the food hoarding behavior of non-deprived mice with that of mice fed the stock pellet diet plus a supplement would be of interest. Wolfe (13) noted that the "genesis of food storing is closely related to diet, but once started it persists though to a lesser extent . . . (13, p. 107)."

On the basis of the above discussion, it may be hypothesized that (a) mice will hoard more when satiation is the initial condition than when deprivation is the initial condition under which hoarding occurs, and, (b) mice fed the stock pellet diet will hoard more than those mice fed the pellets plus a supplement. These hypotheses propose that food deprivation and associated tissue needs inhibit the frequency of hoarding by mice. Experiment I was designed to test these hypotheses.

If food deprivation interferes with hoarding frequency, it is probable that other deprivation also will interfere with the process. On the other hand, "needs" not related to physiological factors (in the same sense as nutritional needs might be related) should not interfere with the frequency of hoarding. A sharp decrease in temperature, for example, should provide sufficient stress to interfere with the frequency of food hoarding behavior. Experiment II provides a tentative answer to this problem. The only previous investigation of the relation of temperature to frequency of hoarding is that by McCleary and Morgan (4) in which it was found that the frequency of hoarding was inversely related to temperature, i.e., as temperature decreased, hoarding frequency increased within wide limits. We have predicted that under stress (low temperature) mice will react somewhat differently than rats.

B. METHOD AND PROCEDURE

1. *Experiment I*

The purpose of this experiment was to compare the hoarding of groups of mice which were (a) deprived of food, (b) fed food pellets *ad libitum*, and (c) fed food pellets plus a dietary supplement.

Seventy-seven male C3H mice, approximately five months old, from the Bucknell University colony were the subjects in this experiment. The mice

were weaned on food pellets and housed in groups of approximately 12 from weaning until four days before the experiment began. As far as could be determined, none of the mice had experienced major food or water deprivation. The litters to which these mice belonged were limited to eight in number by removing the excess young in larger litters shortly following birth. This was done to lessen the competition at the nipple.

Twenty-five of the mice were assigned to Group 1 and 22 to Group 2. Six litters were used and distributed between these two groups, in addition to a control group ($N = 30$).

Four days before the experiment began, all subjects were transferred to the individual living cages. Food pellets and water were provided *ad libitum* to all animals. The individual living cages were arranged in batteries of nine each. The cages and alleys have been described previously (12). On the first test day, the alleys were placed in front of the batteries of cages, the doors raised, and the mice permitted to hoard for 20 minutes. At the end of the trial, the pellets were counted and an adequate supply of food left in each cage until the end of the fifth trial when the animals of Group 1 were given a single pellet while the animals of Group 2 were given a supplement of lettuce and a cube of sugar. Following the tenth trial, Group 1 was supplied food pellets *ad libitum* and the supplement for Group 2 was stopped. The complete plan of the experiment is given in Table 1, showing the details of procedure for Groups 1 and 2, and the control group.

All trials were conducted each morning at eight o'clock. The temperature of the colony room where the experiments were conducted varied between 70°-80° F. Lighting was supplied by a battery of eight, 20-watt, shaded fluorescent lights suspended four feet above the living cages.

2. Experiment II

The purpose of this experiment was to test the hypothesis that those stress conditions which are more or less related to nutritional needs will interfere with the amount of food hoarded. The variable factor (stress) was a sharp variation in environmental temperature. Since no facilities, such as a controlled climatic chamber, were available, a relatively crude substitute was used. Each animal was subjected to a range of temperatures, and thus acted as its own control.

The animals used in the experiment were 30 female C3H mice from the Bucknell colony. They were approximately six months of age when the experiment began.

The mice were transferred to the individual living cages four days prior

TABLE 1
SHOWING THE PLAN OF EXPERIMENT I

Test days	Period	Control ($N=30$)	Group 1 ($N=25$)	Group 2 ($N=22$)
1-5	I	Deprivation: One pellet daily	Satiation: Pellets <i>ad. lib.</i>	Satiation: Pellets <i>ad. lib.</i>
6-10	II	Satiation: Pellets <i>ad. lib.</i>	Deprivation: One pellet daily	Satiation plus supplement: Pellets <i>ad. lib.</i> , lettuce and sugar
11-15	III	Deprivation: One pellet daily	Satiation: Pellets <i>ad. lib.</i>	Satiation: Pellets <i>ad. lib.</i>

to the start of the experiment and were given an adequate supply of food pellets daily throughout the 30 days of the experiment. The mice were housed in a room in which the temperature was varied by manipulating the steam valve, door, and window.

Two recording thermometers were used to record the high temperature and low temperature for the preceding 24-hour period, and to report the temperature at the time of the hoarding trials.

Each morning at eight, the alleys were placed in front of the batteries of the living cages, the temperatures recorded, and the mice were given a 20-minute hoarding trial. Following the trial, all pellets were counted, and as the mice always hoarded under satiation conditions, several pellets were left in each cage. In general, we attempted to give several successive trials at low, medium, and high temperatures. Since mice are extremely susceptible to pneumonia, it was not possible to express them to freezing temperatures.

C. RESULTS AND DISCUSSION

1. *Experiment I*

A graphic presentation of the results of this experiment is given in Figures 1 and 2. This method of presenting the results is considered satisfactory in view of the relatively small numbers of subjects used and the difficulty of applying standard statistical methods to hoarding data (5).

Figure 1 shows the performance of the 25 subjects of Group 1, the group which had experienced no known food deprivation when the hoarding trials began. For purposes of comparison, the performance of the control group which experienced deprivation, satiation, and deprivation respectively for Periods I, II, and III, is included in the figure. The differences in hoarding level between the groups for Periods I and III are large enough to

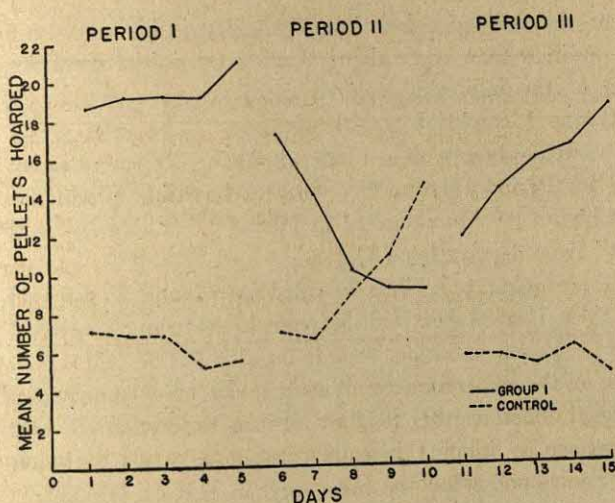


FIGURE 1

THE MEAN NUMBER OF FOOD PELLETS HOARDED DAILY DURING THE THREE TEST PERIODS FOR GROUP 1 AND A CONTROL GROUP

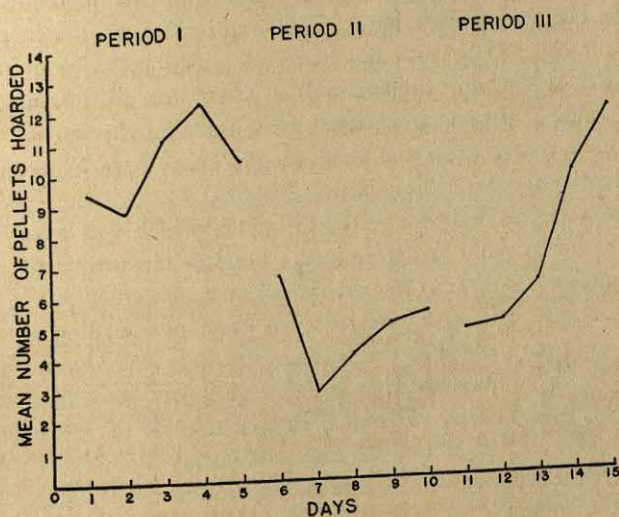


FIGURE 2

THE MEAN NUMBER OF FOOD PELLETS HOARDED DAILY DURING THE THREE TEST PERIODS: (I. SATIATION, II. SATIATION PLUS SUPPLEMENT, AND III. SATIATION) FOR GROUP 2

make it clear that for mice hoarding frequency is higher when satiation is the initial condition than when deprivation is the initial condition. Further confirmation of the inhibiting rôle of deprivation is seen in Period II in which the Group 1 condition was deprivation. The level drops more than 60 per cent, after only five days of deprivation. When satiation conditions are resumed for Period III, the frequency of hoarding climbs daily but does not reach as high a level as occurred in Period I. This may be the result of a "hangover" from deprivation (8).

If the total number of pellets hoarded by Group 1 for each period is considered, there is a 38 per cent decrease in hoarding when deprivation is the condition. When satiation was resumed (Period III) only a 27 per cent increase occurred, indicating that the effects of deprivation influence later hoarding. These results tend to confirm the hypothesis that food deprivation depresses or inhibits hoarding response, while satiation facilitates or enhances the response for the mice used in this experiment.

What we have called deprivation, one food pellet per day, does not constitute a severe deprivation for mice. Mice regularly consume approximately two pellets per day, and we found that mature mice maintained at least constant weight on a single pellet for a 15-day period. It is, therefore, difficult to comprehend why this mild deprivation serves to depress hoarding so severely. In the typical experiment, mild deprivation results in an increase in the amount hoarded by rats and hamsters. The behavior of mice under deprivation has important implications in their use in learning and discrimination studies. In an unpublished maze learning experiment with C3H mice, we found that a preferred food (sugar cube) was a better incentive than food deprivation for either six or 12 hours.

Figure 2 shows the performance of Group 2, which was never subject to deprivation as far as could be determined. Food pellets ordinarily are considered to be an adequate diet for mouse, rat, and hamster colonies. An occasional supplement of lettuce is recommended. In previous hoarding studies, satiation usually has been defined as the feeding of pellets *ad libitum*. The curves in Figure 2 seem to indicate that satiation requires something more than is provided by food pellets. Group 2 carried a total of 1,296 pellets during Period I, 601 during Period II, and 910 during Period III. When the diet was supplemented, hoarding frequency decreased 54 per cent. When the supplement was stopped, the amount hoarded in the following five-day period increased 51 per cent.

Perhaps we should adopt a policy of referring to relative satiation unless it can be ascertained that all known dietary needs are satisfied. Wolfe (13),

in comparing the influence of "adequate" and "inadequate" diets on the amount hoarded by rats, states that:

... The fact that rats raised on a diet sufficient in quality to enable them to reach maturity, reproduce, and appear normal in every way, will readily store food even when non-hungry while other rats given a more complete diet will not indicates that the absence of certain food elements provides distinct motivation for specific behavior (p. 107).

2. Experiment II

In the second experiment, it was predicted that a severe decrease in temperature below that considered normal for a mouse colony would constitute a stress situation which would decrease the hoarding level as does food deprivation. Some of the data obtained from the experiment are shown in Figure 3.

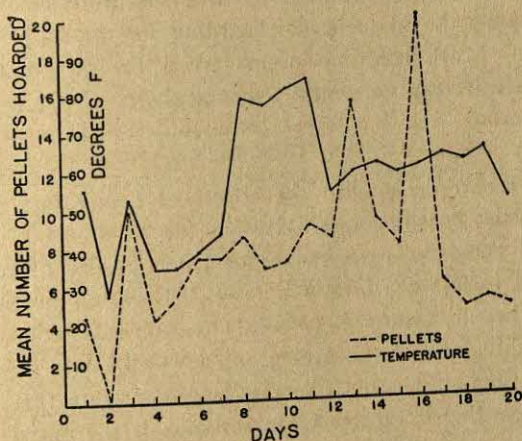


FIGURE 3

THE MEAN NUMBER OF FOOD PELLETS HOARDED DAILY AND LOWEST TEMPERATURE RECORDED DAILY FOR THE 20-DAY TEST PERIOD

These data are based on the records of only 20 mice of the original group since ten of them died on the second test day when the temperature dropped to 28°F.

There appears to be no clear relationship between the lowest temperature for each day and the amount hoarded. Yet, careful scrutiny of the pellet and temperature curves reveals that on 17 of the 20 days the amount hoarded varied directly with the low temperature for the day, i.e., when the temperature increased, the amount hoarded increased and vice versa. This

finding receives further confirmation if the amount hoarded on the nine days (when the temperature was below 59°F.) is compared with the amount hoarded on the 11 days when the temperature was 59°F. or above. Sixty-six per cent of the pellets were hoarded when the temperature was 59°F. or higher, while only 34 per cent of the total was hoarded at temperatures below 59°F.

It will be remembered that the high temperature for each day and the temperature at the time of each morning trial also were recorded. An analysis of these measures with hoarding yielded no additional relevant information. The differences between the high and low temperatures for each day were computed and an inverse relationship between the difference and the amount hoarded was found for 15 of the 20 days. When the size of the difference increased, the amount hoarded decreased, and when the size of the difference decreased, the amount hoarded increased.

The above findings indicate tentatively that temperatures below 59°F. served to inhibit or depress the frequency of hoarding for the subjects of the present experiment. More hoarding occurs when the temperature is nearer the 60°-70°F. normal environmental temperature.

D. SUMMARY

The present report deals with the effects of deprivation, satiation, and stress on food pellet hoarding in C3H mice.

Experiment I compares performance of three groups of mice for three hoarding periods, as follows: Group 1. ($N=25$) was on a satiation, deprivation, satiation schedule, Group 2. ($N=22$) was on a satiation, satiation plus supplement schedule, and the control group ($N=30$) was on a deprivation, satiation, deprivation schedule. The results demonstrate that the frequency of food pellet hoarding is greater when satiation is the initial testing condition. Deprivation thus seems to play an inhibiting rôle. The relationship between satiation and hoarding frequency, especially in regard to a lettuce and sugar supplemented diet, is complex. The food supplement period yielded a major decrease in the total number of pellets hoarded from 1,296 pellets during the five days of satiation hoarding to 601 during a similar period of supplemental feeding. A return to the previous state of satiation feeding (without supplement) resulted in an increase in hoarding frequency.

Experiment 2 describes the effect of exposure to cold on hoarding. Some relationship between the number of pellets hoarded and the lowest temperature recorded for each seems to be present, i.e., as the lowest daily temperature increased, the number of food pellets hoarded increased. It is concluded, tentatively that temperatures below approximately 59°F. serve

to inhibit the frequency of hoarding. Greater hoarding frequency occurs when the temperature is approximately 60°-70° F.

REFERENCES

1. BINDRA, D. The nature of motivation for hoarding food. *J. Comp. & Physiol. Psychol.*, 1948, **41**, 211-218.
2. COMMITTEE ON FOOD HABITS, NAT. RES. COUNCIL. Manual for the Study of Food habits. *Bull. Nat. Res. Council*, 1945.
3. LICKLIDER, L. C., & LICKLIDER, J. C. R. Observations on the hoarding behavior of rats. *J. Comp. & Physiol. Psychol.*, 1950, **43**, 129-134.
4. MCCLEARY, R. A., & MORGAN, C. T. Food hoarding in rats as a function of environmental temperature. *J. Comp. Psychol.*, 1946, **39**, 371-378.
5. MORGAN, C. T. The statistical treatment of hoarding data. *J. Comp. Psychol.*, 1945, **38**, 247-256.
6. ———. The hoarding instinct. *Psychol. Rev.*, 1947, **54**, 335-341.
7. MORGAN, C. T., STELLAR, E., & JOHNSON, O. Food deprivation and hoarding in rats. *J. Comp. Psychol.*, 1943, **35**, 275-295.
8. ROSS, S., SMITH, W. I., & NIENSTEDT, C. W., JR. The hoarding of non-relevant material by the white rat. *J. Comp. & Physiol. Psychol.*, 1950, **43**, 217-225.
9. SMITH, W. I., & ROSS, S. The effect of continued food deprivation on hoarding in the albino rat. *J. Genet. Psychol.*, 1950, **77**, 117-121.
10. ———. Hoarding behavior in the golden hamster (*mesocricetus auratus auratus*). *J. Genet. Psychol.*, 1950, **77**, 211-215.
11. ———. A note on food hoarding by adrenalectomized hamsters. (In preparation.)
12. ———. The hoarding behavior of the mouse: I. The rôle of previous feeding. *J. Genet. Psychol.*, 1953, **82**, 279-297.
13. WOLFE, J. B. An exploratory study in food storing in rats. *J. Comp. Psychol.*, 1939, **28**, 97-101.

Department of Psychology
University of Maryland
College Park, Maryland

Department of Psychology
Bucknell University
Lewisburg, Pennsylvania

THE HOARDING BEHAVIOR OF THE MOUSE: III. THE STORING OF "NON-RELEVANT" MATERIAL*

Departments of Psychology, Bucknell University and University of Maryland

WENDELL SMITH AND SHERMAN ROSS¹

A. INTRODUCTION

In previous papers on hoarding by mice, the rôles of pellet experience (10), deprivation, satiation, diet, and temperature (11) have been considered. The mouse was found to be an adept hoarder under satiation conditions. In fact, mice which had never been deprived of food pellets hoarded at a higher frequency than those subjected to food deprivation. In this paper, we shall be concerned with the frequency of the hoarding of so-called "*non-relevant*" materials.

The term, *non-relevant*, requires some explanation. One might ask, non-relevant to what? As used in previous investigations of hoarding by Ross, Smith, and Nienstedt (8), Morgan (7) citing an unpublished study by Johnson, Miller and Viek (6), the term has implied non-relevance to the "known" needs of the animal. If a thirsty animal hoards food pellets, the pellets may be said to be non-relevant to the animal's present needs. If male rats hoard material which can be used for nests to cages which already contain ample supplies of bedding, the material carried may be termed non-relevant. There may be some disagreement with this concept of relevance. The basic factor involved, however, is the concept of drive and the related factor of need. If wooden blocks are carried into the cages, do the wooden blocks, or does the activity of carrying, result in need reduction? If hoarding is motivated, it must be motivated by something, and it must satisfy this "something." If an underlying need is assumed, and if a non-relevant material, such as wooden blocks, is hoarded, it must be some aspect of the hoarding behavior which acts to reduce the need. On the other hand, if non-relevant material is not hoarded, then we could assume that the goal of hoarding is the acquisition of material relevant to, or closely associated with, the needs of the animal.

There is little evidence available on this point inasmuch as the majority of hoarding studies have utilized food as the to-be-hoarded object, and, in

*Received in the Editorial Office on May 23, 1951.

¹This study was supported in part by a grant-in-aid from Bucknell University.

addition, have usually utilized food deprived subjects. In a few cases, satiated subjects have been used with food as the hoarding object. Licklider and Licklider (2), for example, report a study in which satiated rats, weaned to food pellets, hoarded plain and foil-wrapped pellets, salted peanuts, peanuts in shells, and chocolates. These writers propose that perceptual (visual, tactual, taste) factors should be emphasized in a theory of hoarding. The hoarding of these various food substances may be a generalization from previous pellet experience, and if peanuts and chocolates are more preferred foods than pellets, the satiation hoarding of larger amounts of these foods than of pellets may be attributed to this latter factor as well. The higher frequency of the hoarding of foil covered pellets indicates a preference for these objects over plain pellets. As olfactory cues for these objects are present, as well as possible taste cues from gnawing through the foil, the storing of these objects is not irrelevant to the past needs of the organism.

In a very stimulating paper on the hoarding of various materials by the rat, Marx (3) reports the hoarding of 10 wooden blocks by one satiated rat, and the hoarding of nine blocks by a second rat following food deprivation. Neither of these rats had had previous experience in hoarding alleys, therefore Marx attributes this behavior to the possible "... operation of stimulus generalization, from the feeding responses made to real good pellets in the home cage." He says further that the "... generalization probably occurred because of the absence of the more effective stimulation from the food pellets" (3, p. 303). Miller and Vick (6), Johnson (7), and Ross et. al. (8) have all reported the non-hoarding of wooden blocks, although in the Ross study several rats carried blocks out of the cages and deposited them at the end of the runway.

All this would seem to indicate that what is hoarded to the home cage has some relevance to the animal's needs. The only exceptions to this point are the two rats in the Marx study, and anecdotal reports of the hoarding of irrelevant materials by a few rodent groups such as hamsters (5). If materials which appear to be irrelevant are hoarded freely, we might ask what needs are satisfied. If we assume that hoarding is learned, (3, 4) what is the reinforcement for carrying and storing the material? In the present study the hoarding by mice of wooden blocks and cotton packs was investigated.

B. METHOD AND PROCEDURE

The purpose of this experiment was to determine the amount of "non-relevant" material which mice would hoard following previous hoarding experience with food pellets.

The subjects for the study were 18 male C3H mice from the Bucknell colony approximately six months old. The mice were weaned on food pellets and housed in groups until selected for the experiment. As far as could be determined, these mice had never experienced food or water deprivation. The groups in which they were placed were small, and food and water were provided at all times. Lettuce was supplied to the colony once each week.

Four days prior to the beginning of the experiment, the mice were transferred in individual living cages and maintained on a satiation schedule throughout the 25 test days of the experiment. The rationale for the four-day adaptation period and a description of the cages and hoarding alleys have been presented elsewhere (10). Wood shavings were present at all times in the group cages as well as in the hoarding cages.

A schedule of the experiment is given in Table 1. It will be noted that

TABLE 1
SHOWING THE SCHEDULE OF THE EXPERIMENT

Period	Test days	Material to be hoarded
1.	1-5	Food pellets
2.	6-10	Cotton pellets (dry)
3.	11-15	Food pellets
4.	16-20	Cotton pellets (wet)
5.	21-25	Food pellets

the mice were tested first with food pellets in the hoarding alley before the "non-relevant" dry cotton dental plugs were presented in the alley in Period 2. For Period 4 the cotton plugs were immersed in water for five minutes before being placed in the hoarding alleys.

On the first day of the experiment, the alleys were placed before the cages, the doors were raised, and the mice were given a 20-minute hoarding trial. At the end of the trial, the mice were returned to the cages, the alleys removed and washed, and the pellets of food or cotton which had been carried into the cages were removed. A supply of food pellets was left in each cage following each trial.

C. RESULTS AND DISCUSSION

The mice in this experiment were never subjected to food or water deprivation, and particular care was taken to insure ample dry bedding. We shall be interested then, in the hoarding performance of these animals, particularly for Periods 2 and 4, the periods in which respectively, dry and wet cotton pellets were the hoarding objects. A graphic presentation of the mean number of pellets hoarded by the group for each of the 25 test days is given in Figure 1.

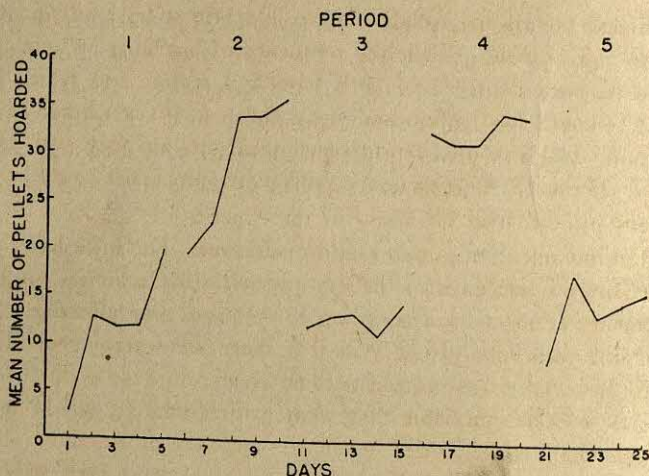


FIGURE 1

THE MEAN NUMBER OF PELLETS HOARDED BY THE GROUP ($N=18$) FOR EACH DAY OF THE EXPERIMENT: PERIODS 1, 3, AND 5 REPRESENT FOOD PELLET HOARDING, WHILE PERIOD 2 IS DRY COTTON PELLET HOARDING AND PERIOD 4 IS WET COTTON PELLET HOARDING

There are several points of interest in this figure. First it should be noted that the hoarding of food pellets started at a low level, but increased rapidly during the remainder of Period 1. The frequency of hoarding food pellets in Periods 3 and 5 appears to be at about the same level as in Period 1, excepting the first day of the experiment. Secondly, the frequency of hoarding the dry cotton plugs is considerably above that of the food pellets, and the frequency of hoarding the wet cotton packs is the highest of all. The mean number of pellets hoarded by the group for each of the five periods is as follows: Period 1: 11.7; Period 2: 28.92, Period 3: 12.81, Period 4: 32.49, and Period 5: 14.09. These data help to support the interpretation of the curves in Figure 1.

The astonishing finding in these results is the significant increase in the number of pellets hoarded when cotton packs were substituted in the alleys for the food pellets. The means for the cotton packs are more than double those for the food pellets. Why are the cotton packs hoarded at this high rate, in fact why are the dry cotton packs hoarded at all? Marx (3) presents some data on the hoarding of dry cotton pellets and wooden blocks by six rats which previously had been given a choice of food pellets and wet cotton packs following food or water deprivation. When satiated and given dry cotton pellets and wooden blocks, the rats hoarded "... an average of about

four cotton pellets and one wooden block, and left at least an equal number in the alleys" (3, p. 304). This low frequency of hoarding by animals which previously had been deprived of both food and water, and which had had opportunity to hoard food pellets and wet cotton packs is attributed to stimulus generalization. In terms of a reinforcement learning analysis, such as the one preferred by Marx (3, 4), secondary reinforcing properties could be assigned the alley and/or the food bin.

These data on mice are not so easy to interpret. The mice had never had previous experience with cotton pellets nor with any material similar to it. The attributing of the cotton hoarding to stimulus generalization from food pellet hoarding seems unacceptable in that there are several gross differences in the stimulus qualities of cotton pellets when compared with food pellets. Furthermore, it is not probable that a far greater amount would be hoarded on this basis than was hoarded when the primary or standard object, the food pellet, was the hoarded object. There could have been secondary reinforcing properties from the alley and food bin, if the assumption of reinforcement learning of hoarding is valid, but this influence alone does not appear to provide a satisfactory explanation.

The hoarding of wooden blocks, similar in size and shape, to food pellets, perhaps is not non-relevant hoarding since the operation of stimulus generalization is more plausible in this instance. From the literature, we learn that only rats actually have carried and stored more than two or three blocks in the home cage (3). Will mice hoard wooden blocks when they are presented with a choice of blocks and food pellets following 15 days of food pellet hoarding experience? Some data are available on this point.

Using a group of 20 C3H male mice which had completed 15 hoarding trials for another study, we placed 50 food pellets and 10 wooden blocks in each alley, and gave the group 20-minute trials on three successive days. Even though we had weighted the situation in favor of pellet hoarding, 30.5, 44.5, and 57.0 per cent of the blocks were hoarded on successive days compared with 44.9, 48.0 and 47.3 per cent of the pellets. Of the 60 scores obtained for blocks, only seven were zero, and only one of the 20 mice failed to hoard any blocks, accounting for three of the seven zero scores. The median number of blocks hoarded was five. Thus, mice hoarded wooden blocks more frequently than did rats. This may be the result of a lower ability to discriminate pellets from blocks (3, 6). However, as the number of blocks hoarded increased approximately 27 per cent from Trial 1 to Trial 3, one might get the impression that discriminative ability became progressively poorer or that mice were either learning to or developing preference for, wooden blocks.

Earlier the question of need in relation to the object hoarded was raised. What need was reduced by the carrying and storing of the dry cotton pellets? Adequate bedding in the form of wood shavings was available, the temperature of the colony room never fell below 70°F, and the subjects were males.

It is, of course, possible that nesting material was the "need" underlying this behavior. The experimenter was always present in the room during all hoarding trials. For Period 2, when dry cotton pellets were being hoarded, some mice were observed, even on the first day, pulling one or two of the cotton pellets apart until a fairly large ball of cotton was present in the cage. This was not the typical pattern, and it never occurred until the second half of the hoarding trial. None of the mice which behaved in this manner were found in or on the "nest" of cotton. This behavior did not occur when the wet pellets were presented in Period 4. Miller and Vick, discussing the relationship of hoarding to nesting, have the following to say:

... Nesting behavior, like hoarding, is essentially the accumulation of objects into the cage. Their similarity easily suggests the possibility that they are somehow related instincts. If the two patterns are related, however, one would not expect the rats to discriminate between pellets and blocks, since any object which could be carried should satisfy the behavior pattern, and blocks are as good nesting material as pellets. . . it was found, however, that the rats do discriminate between pellets and blocks, and carry only the pellets. In view of this fact, and also of the general importance of deprivation to the appearance of the hoarding pattern, one is forced to doubt the validity of the hypothesis that hoarding is a modification of the nesting instinct (6, p. 230).

In the present study, mice did carry wooden blocks and dry cotton pellets, and deprivation was not of "general importance." Therefore, following the Miller and Vick discussion, hoarding still might be a "modification of the nesting instinct."

An impression of a greater amount of activity in Periods 2 and 4, raised the question of the place or function of the activity level in hoarding. Nearly every writer on this subject has noted the presence of this factor, but apparently no one has controlled or varied activity level in relation to hoarding frequency. The subjects in the studies which have been reported have all been confined in relatively small cages during the experiment. When so-called satiated animals exhibit a high activity level, some physiological "deficit" or "need" would appear to have been overlooked in designing the experiment. Scott (9) found no hoarding of food pellets by mice which were housed in a large multiple escape pen, yet in the same strain of mice, we have found a high level of hoarding of food, dry and wet cotton, and wooden pellets by satiated

subjects. At least two factors may be responsible for this difference between the mice in Scott's study and the mice used in the several hoarding studies conducted by the writers. First the mice used in Scott's multiple escape pen had more opportunity for activity, i.e., they were less confined; and, secondly, Scott's mice may have been more "secure" (less "shy") in the sense in which Bindra (1) has applied these terms to the hoarding situation. It should be possible to conduct a study which would throw some light on the rôle played by these two factors. Unless the motivation for hoarding dry cotton pellets at a level above that of the food pellets is discovered, it would appear that mice, and perhaps other species, will carry and store objects which are actually non-relevant.

D. SUMMARY

The present experiment deals with the hoarding of so-called "non-relevant" objects by mice. A group of 18 male C3H mice approximately six months of age were tested for the hoarding of food pellets and dry and wet cotton packs. The mice had never been subjected to food or water deprivation either previous to or during the experimental period of 25 days.

It was found that the subjects hoarded both the dry and wet cotton pellets. In addition, the mean number of cotton packs hoarded was far greater than the number of food pellets hoarded during each of three test periods.

The significance of the high rate of cotton pack hoarding by these satiated mice is discussed in terms of previous hypotheses as to the nature of the hoarding of "non-relevant" objects.

REFERENCES

1. BINDRA, D. What makes rats hoard? *J. Comp. & Physiol. Psychol.*, 1948, **41**, 397-402.
2. LICKLIDER, L. C., & LICKLIDER, J. C. R. Observations on the hoarding behavior of rats. *J. Comp. & Physiol. Psychol.*, 1950, **43**, 129-134.
3. MARK, H. Experimental analysis of the hoarding habit in the rat. I. Preliminary observations. *J. Comp. & Physiol. Psychol.*, 1950, **43**, 295-308.
4. ———. A stimulus-response analysis of the hoarding habit in the rat. *Psychol. Rev.*, 1950, **57**, 80-93.
5. McLELLAN, M. The Syrian hamster. *Nat. Mag.*, 1942, **35**.
6. MILLER, G. A., & VIEK, P. An analysis of the rat's response to unfamiliar aspects of the hoarding situation. *J. Comp. Psychol.*, 1944, **37**, 221-231.
7. MORGAN, C. T. The hoarding instinct. *Psychol. Rev.*, 1947, **54**, 333-341.
8. ROSS, S., SMITH, W. L., & NIENSTEDT, C. W. JR. The hoarding of non-relevant material by the white rat. *J. Comp. & Physiol. Psychol.*, 1950, **43**, 217-225.
9. SCOTT, J. P. Social behavior, range, and territoriality in domestic mice. *Proc. Ind. Acad. Sci.*, 1944, **53**, 188-195.

10. SMITH, W., & ROSS, S. The hoarding behavior of the mouse: I. The rôle of previous feeding experience. *J. Genet. Psychol.*, 1953, **82**, 279-297.
11. ———. The hoarding behavior of the mouse: II. The rôle of deprivation, satiation, and stress. *J. Genet. Psychol.*, 1953, **82**, 299-307.

Department of Psychology
Bucknell University
Lewisburg, Pennsylvania

Department of Psychology
University of Maryland
College Park, Maryland

BOOKS

The *Journal of Genetic Psychology*, the *Journal of General Psychology*, and the *Journal of Social Psychology*, will buy competent reviews at not less than \$2 per printed page and not more than \$3 per printed page, but not more than \$15.00 for a single review.

Conditions. Only those books that are listed below in this section are eligible for such reviews. In general, any book so listed contains one or more of the following traits: (a) Makes an important theoretical contribution; (b) consists largely of original experimental research; (c) has a creative or revolutionary influence in some special field or the entire field of psychology; (d) presents important techniques.

The books are listed approximately in order of receipt, and cover a period of not more than three years. A reviewer must possess the Ph.D. degree or its equal in training and experience.

Procedure. If among the books listed below there is one that seems important to you, you are invited to write a review of that book. It is not necessary to make arrangements with the Editor. Just send in your review. It does not matter if the book in question has been reviewed before.

1952

- STONE, C. P., Ed. *Annual Review of Psychology*, Vol. 3. Stanford, Calif.: Annual Reviews, 1952. Pp. 462.
- HOCH, P. H. & ZUBIN, J., Eds. *Relation of Psychological Tests to Psychiatry*. New York: Grune & Stratton, 1952. Pp. 301.
- CATTELL, R. B. *Factor Analysis*. New York: Harper, 1952. Pp. 462.
- FOERSTER, H. VON, Ed. *Cybernetics*. New York: Macy Foundation, 1952. Pp. 240.
- BENDER, L. *Child Psychiatric Techniques*. Springfield, Ill.: 1952. Pp. 335.
- METTLER, F. A., Ed. *Psychosurgical Problems*. New York: Blakiston, 1952. Pp. 357.
- BENDA, C. E. *Developmental Disorders of Mentation and Cerebral Palsies*. New York: Grune & Stratton, 1952. Pp. 565.
- KINGET, G. M. *The Drawing-Completion Test*. New York: Grune & Stratton, 1952. Pp. 238.
- WOLFF, C. *The Hand in Psychological Diagnosis*. New York: Philosophical Library, 1952. Pp. 218.
- WOLFF, W. *The Dream: Mirror of Conscience*. New York: Grune & Stratton, 1952. Pp. 348.
- GORLOW, L., HOCH, E. L., & TELSCHOW, E. F. *The Nature of Nondirective Group Psychotherapy*. New York: Teach. Coll., Columbia Univ., 1952. Pp. 143.
- AUSUBEL, D. P. *Ego Development and the Personality Disorders*. New York: Grune & Stratton, 1952. Pp. 564.
- BYCHOWSKI, G., & DESPERT, J. L. *Specialized Techniques in Psychotherapy*. New York: Basic Books, 1952. Pp. 371.
- WHITE, R. W. *Lives in Progress: A Study of the Natural Growth of Personality*. New York: Dryden, 1952. Pp. 376.
- KARN, H. W., & GILMER, B. VON H. *Readings in Industrial and Business Psychology*. New York: McGraw-Hill, 1952. Pp. 476.
- ROBACK, A. A. *History of American Psychology*. New York: Library Pub., 1952. Pp. 426.
- KROEBER, A. L. *The Nature of Culture*. Chicago: Univ. Chicago Press, 1952. Pp. 437.
- FERGUSON, L. W. *Personality Measurement*. New York: McGraw-Hill, 1952. Pp. 457.

- BECK, S. J. *Rorschach's Test: III. Advances in Interpretation*. New York: Grune & Stratton, 1952. Pp. 301.
- VINACKE, W. E. *The Psychology of Thinking*. New York: McGraw-Hill, 1952. Pp. 392.
- KUHLEN, R. G., & THOMPSON, G. G., *Eds.* *Psychological Studies of Human Development*. New York: Appleton, Century, Crofts, 1952. Pp. 533.
- SNYDER, F. W., & PRONKO, N. H. *Vision with Spatial Inversion*. Wichita: Univ. Wichita Press, 1952. Pp. 144.
- JERSILD, A. T. *In Search of Self*. New York: Teach. Coll., Columbia Univ., 1952. Pp. 141.
- HIRSH, O. J. *The Measurement of Hearing*. New York: McGraw-Hill, 1952. Pp. 364.
- ASHBY, W. R. *Design for a Brain*. New York: Wiley, 1952. Pp. 260.
- SHAFFER, G. W. *Fundamental Concepts in Clinical Psychology*. New York: McGraw-Hill, 1952. Pp. 540.
- HAYEK, F. A. *The Sensory Order*. Chicago: Univ. Chicago Press, 1952. Pp. 209.
- AMES, L. B., LEARNED, J., METRAUX, R. W., & WALKER, R. N. *Child Rorschach Responses*. New York: Harper, 1952. Pp. 310.

1953

- KLUCKHORN, C., MURRAY, H. A., & SCHNEIDER, D. M., *Eds.* *Personality*. New York: Knopf, 1953. Pp. 701.
- SKINNER, B. F. *Science and Human Behavior*. New York: Macmillan, 1953. Pp. 461.
- STONE, C. P., & TAYLOR, D. W., *Eds.* *Annual Review of Psychology*. Stanford: Annual Reviews, 1953. Pp. 485.
- BURGESS, E. W., & WALLIN, P. *Engagement and Marriage*. New York: Lippincott, 1953. Pp. 819.
- POWDERMAKER, F. B., & FRANK, J. D. *Group Psychotherapy*. Cambridge: Harvard Univ. Press, 1953. Pp. 615.
- PHILLIPS, L., & SMITH, J. G. *Rorschach Interpretation: Advanced Technique*. New York: Grune & Stratton, 1953. Pp. 385.
- MURPHY, G. *In the Minds of Men*. New York: Basic Books, 1953. Pp. 306.
- BLUM, G. S. *Psychoanalytic Theories of Personality*. New York: McGraw-Hill, 1953. Pp. 219.
- BROWER, D., & ABT, L. E., *Eds.* *Progress in Clinical Psychology*. New York: Grune & Stratton, 1953. Pp. 564.
- GELDARD, F. A. *The Human Senses*. New York: Wiley, 1953. Pp. 365.
- CABOT, H., & KAHL, J. A. *Human Relations*. Cambridge: Harvard Univ. Press, 1953. Pp. 333 & 273.
- SHERIF, M., & SHERIF, C. W. *Groups in Harmony and Tension*. New York: Harper, 1953. Pp. 316.
- WHITING, J. W. M., & CHILD, I. L. *Child Training and Personality*. New Haven, Conn.: Yale Univ. Press, 1953. Pp. 353.

CRITICAL REVIEWS OF RECENT BOOKS

The Journal of Genetic Psychology, 1953, 82, 319-321.

(Greene, E. B. *Measurements of Human Behavior*. New York: Odyssey Press, 1952. Pp. 790.)

REVIEWED BY ROBERT M. ALLEN

This book, originally published in 1941, has found general acceptance either as the main or supplementary text for courses in Tests and Measurements. This reviewer has considered Greene's volume as a less expensive substitute for Buross' *Mental Measurements Yearbook*. The 1952 revision is a more complete presentation of the techniques for measuring all aspects of human behavior. The expansion of the scope of information covered by the revised text includes the pictures of leaders in tests and measurements. The insertion of 31 pictures is an increase of 28 over the first edition—indicating, no doubt, that the three *T*'s (Terman, Thorndike, and Thurstone) are now in good company.

The physical changes are few: the new edition lists the names of tests appropriately classified according to type of process measured, along with names and addresses of test publishers. The bibliography is more current. An obvious omission is the Glossary which formed part of the Subject Index in the first edition. The student, for whom this book is designed, is deprived of a source of review as well as definition.

The content changes are many. The author retained the basic material as the foundation for presenting more complete information in the light of research data accumulated during the 10 or 11 years between editions. Judging from the new placement of the contents (or chapters) it is quite apparent that the author benefited from the intervening years of experience regarding the optimum order of teaching basic information in a course on the theory, construction, and use of tests. If the present volume is followed as printed, the student is very early apprised of sources of information regarding research and ethical considerations in test use and distribution. It is essential that neophytes do learn the problems that the professional psychologist must face and solve.

The first part of the revised text is devoted to theoretical and practical material necessary for a solid foundation in testing. This could prove to be most useful to the student entering the field of psychological and/or educational testing. This volume is surpassed only by the *Yearbook* as a catalogue and

source of vital information on various paper and pencil, performance, and individual psychological tests. This alone would justify its inclusion in the applied psychologist's and educational tester's book shelf.

A history of the testing movement could not overlook the advances in test construction and use achieved during World War II. Greene's new chapter, *Military Development of Tests and Ratings*, brings the reader up to date on the rôle of tests in the tremendous task of assigning 12,000,000 men and women to some 800 to 900 different military occupational jobs ranging from combat infantry-man to aviation cadet. The carryover of this experience into a civilian setting is also discussed.

Part two of the revised edition brings together the statistical concepts essential for teaching and learning the quantitative aspects of test construction on the one hand, and the evaluation of the usefulness of a test as a measuring instrument on the other. In courses designed to omit the teaching of statistics the instructor could easily assign this section for informational reading without losing the flavor of the entire book. The presentation of the statistical concepts is quite palatable to students in view of the emphasis on the applications rather than the logical development of the various statistical techniques. This portion could have been strengthened if the author had differentiated the rôles of descriptive and sampling statistics as related to the problems of test construction. The chapter on Factor Analyses is too simplified and might leave the naïve reader with the impression that factor analysis is a pure statistical process. It is this reviewer's opinion that students should be made aware of the compulsion of factor analysts to divide an orange into its component parts—orange peel, orange pulp, and orange juice—three constituents which separately could not possibly be mistaken for the whole, colored, edible citrus fruit.

The third part of this book, *Dynamic Patterns*, is most valuable since it represents the most significant improvement over the original edition. It is a cursory survey of personality theories and related methods of appraisal. This attempt to cover a wide field of study falls far short of what the student might be expected to learn in this area. In the 22 pages of Chapter XV, *Personality: Dynamic Theory and Structure*, there are discussions of physical growth, learning, physical bases for personality appraisal, psycho-analytical theory, and several other topics. If the purpose is to create curiosity, whet the appetite, and stimulate further study, it is adequately accomplished for the serious student. On the other hand, it may give a smattering of ignorance to the student who is willing to less than half-learn, especially Freudian psychology. The rôle of the projective method enjoys a much wider

and intensive treatment in this revision. In the original edition these devices were allotted a total of 15 pages. Today their importance is reflected in the four chapters devoted to drawing, painting, and handwriting; stories and fantasies; play and drama; and Rorschach techniques. While the theoretical rationale is kept to a minimum, there is a surprising amount of material on each of these instruments of personality appraisal (supported by a wide sampling of research reports) which imparts excellent ideas regarding the areas of application for these tests. At best, the coverage is survey in nature; it could not be otherwise in a book such as this one. There is ample reference material so that the reader is not left without additional sources for further study and reading. This third part of the book reflects the *zeitgeist*—it has definitely “gone clinical.”

The revision suffers somewhat from the omission of the earlier section on the persistent problems of the measurement of human behavior. Some of this material is more appropriately placed in other portions, but the student should be sensitized to the problems of practice effect, scaling of units of measurement, and the issues related to the appraisal of physical and psychological growth of the organism. This, incidentally, could introduce the student to the *MA-CA* controversy and the validity of the assumptions behind the *IQ*. The Glossary should also have been included in the new volume.

An overall view of the revised edition reveals a decided improvement of this popular and standard textbook. It is recommended for continued use in undergraduate courses in Tests and Measurements.

Department of Psychology
University of Miami
Coral Gables (University Branch) 46
Florida

BOOKS RECENTLY RECEIVED

(There will always be two pages of book titles, listed in the order of receipt, i.e., the most recently received books will be found at the end of the list.)

- MUKERJEE, R. *The Dynamics of Morals*. London: Macmillan, (Undated). Pp. 530.
- METTLER, F. A., Ed. *Psychological Problems*. New York: Blakiston, 1952. Pp. 357.
- SCHIEDLINGER, S. *Psychoanalysis and Group Behavior*. New York: Norton, 1952. Pp. 245.
- GRAY, J. S. *Psychology in Industry*. New York: McGraw-Hill, 1952. Pp. 401.
- LEE, A. M. *How to Understand Propaganda*. New York: Rinehart, 1952. Pp. 281.
- CUSSLER, M., & DE GIVE, M. L. *'Twixt the Cup and the Lip*. New York: Twayne, 1952. Pp. 262.
- CUSTANCE, J. *Wisdom, Madness, and Folly*. New York: Pellegrini & Cudahy, 1952. Pp. 254.
- BERRIEN, F. K. *Practical Psychology*. New York: Macmillan, 1952. Pp. 640.
- GORLOW, L., HOCH, E. L., & TELSCHOW, E. F. *The Nature of Nondirective Group Psychotherapy*. New York: Teach. Coll., Columbia Univ., 1952. Pp. 143.
- WHITE, R. W. *Lives in Progress: A Study of the Natural Growth of Personality*. New York: Dryden, 1952. Pp. 376.
- BROWER, E. O. *Should Fatherhood Begin at Forty?* Philadelphia: Brower, 1952. Pp. 240.
- KARN, H. W., & GILMER, B. VON H., Eds. *Readings in Industrial and Business Psychology*. New York: McGraw-Hill, 1952. Pp. 476.
- BERGER, M. *Equality by Statute*. New York: Columbia Univ. Press, 1952. Pp. 238.
- ROBACK, A. A. *History of American Psychology*. New York: Library Pub., 1952. Pp. 426.
- HYNDMAN, O. R. *The Origin of Life and the Evolution of Living Things*. New York: Philosophical Library, 1952. Pp. 648.
- RUNES, D. D. *Of God, the Devil, and the Jews*. New York: Philosophical Library, 1952. Pp. 186.
- AUSUBEL, D. P. *Ego Development and the Personality Disorders*. New York: Grune & Stratton, 1952. Pp. 564.
- KERR, M. *Personality and Conflict in Jamaica*. Liverpool: Univ. Liverpool Press, 1952. Pp. 221.
- BYCHOWSKI, G., & DESPERT, J. L., Eds. *Specialized Techniques in Psychotherapy*. New York: Basic Books, 1952. Pp. 371.
- SHOSTROM, E. L., & BRAMMER, L. M. *The Dynamics of the Counseling Process*. New York: McGraw-Hill, 1952. Pp. 213.
- KORNITZER, M. *Child Adoption in the Modern World*. New York: Philosophical Library, 1952. Pp. 403.
- RAVEN, J. C. *Human Nature*. London: Lewis, 1952. Pp. 226.
- MONTAGU, A. *Darwin: Competition and Cooperation*. New York: Schuman, 1952. Pp. 148.
- BARLOW, F. *Mental Prodigies*. New York: Philosophical Library, 1952. Pp. 256.
- JERSILD, A. T. *In Search of the Self*. New York: Teach. Coll., Columbia Univ., 1952. Pp. 141.
- SNYDER, F. W., & PRONKO, N. H. *Vision with Spatial Inversion*. Wichita: Univ. Wichita Press, 1952. Pp. 144.
- EISLER, R. *Man into Wolf*. New York: Philosophical Library, 1952. Pp. 286.
- GUILFORD, J. P. *General Psychology*. New York: Van Nostrand, 1952. Pp. 587.
- KUHLEN, R. G., & THOMPSON, G. G. *Psychological Studies of Human Development*. New York: Appleton, Century, Crofts, 1952. Pp. 533.

- VINACKE, W.E. *The Psychology of Thinking*. New York: McGraw-Hill, 1952. Pp. 392.
- FERGUSON, L.W. *Personality Measurement*. New York: McGraw-Hill, 1952. Pp. 457.
- SHOCK, N.W., *Ed.* *Problems of Aging*. New York: Macy, 1952. Pp. 138.
- McKILLOP, A.S. *The Relationship Between the Reader's Attitude and Certain Types of Reading Response*. New York: Teach. Coll., Columbia Univ., 1952. Pp. 101.
- BECK, S.J. *Rorschach's Test: III. Advances in Interpretation*. New York: Grune & Stratton, 1952. Pp. 301.
- SMILEY, M.B. *Intergroup Education and the American College*. New York: Teach. Coll., Columbia Univ., 1952. Pp. 212.
- GRIFFITHS, W. *Behavior Difficulties of Children*. Minneapolis: Univ. Minnesota Press, 1952. Pp. 116.
- KROEBER, A.L. *The Nature of Culture*. Chicago: Univ. Chicago Press, 1952. Pp. 437.
- HIRSH, J. *Alcohol Education*. New York: Schuman, 1952. Pp. 107.
- ASHBY, W.R. *Design for a Brain*. New York: Wiley, 1952. Pp. 260.
- HIRSH, I.J. *The Measurement of Hearing*. New York: McGraw-Hill, 1952. Pp. 364.
- LEVINSON, A. *The Mentally Retarded Child*. New York: Day, 1952. Pp. 190.
- PERLS, F., HEFFERLINE, R. F., & GOODMAN, P. *Gestalt Therapy*. New York: Julian, 1952. Pp. 466.
- SHAFFER, G. W., & LAZARUS, R. S. *Fundamental Concepts in Clinical Psychology*. New York: McGraw-Hill, 1952. Pp. 540.
- PALMER, H. *The Philosophy of Psychiatry*. New York: Philosophical Library, 1952. Pp. 70.
- HAYEK, F. A. *The Sensory Order*. Chicago: Univ. Press, 1952. Pp. 209.
- POWYS, J. C. *The Inmates*. New York: Philosophical Library, 1952. Pp. 318.
- BERNARD, H. W. *Mental Hygiene for Classroom Teachers*. New York: McGraw-Hill, 1952. Pp. 472.
- WOLFLE, D., *et al.* *Improving Undergraduate Instruction in Psychology*. New York: Macmillan, 1952. Pp. 60.
- JUNGE, W. *Bolahun: An African Adventure*. New York: Putnam, 1952. Pp. 248.
- ITTETSON, W. H. *The Ames Demonstrations in Perception*. Princeton: Princeton Univ. Press, 1952. Pp. 88.
- MUNDT, E. *Art, Form, and Civilization*. Berkeley: Univ. California Press, 1952. Pp. 246.
- SZONDI, L. *Experimental Diagnostics of Drives*. (Trans. by G. Aull.) New York: Grune & Stratton, 1952. Pp. 220.
- DEESE, J. *The Psychology of Learning*. New York: McGraw-Hill, 1952. Pp. 398.
- DE GRAZIA, S. *Errors of Psychotherapy*. New York: Doubleday, 1952. Pp. 288.
- BISNO, H. *The Philosophy of Social Work*. Washington: Public Affairs Press, 1952. Pp. 143.
- BELLAK, L., *Ed.* *Psychology of Physical Illness*. New York: Grune & Stratton, 1952. Pp. 243.
- HALL, B. A., *et al.* *Psychiatric Aide Education*. New York: Grune & Stratton, 1952. Pp. 168.
- BROWER, D., & ABT, L. E., *Eds.* *Progress in Clinical Psychology*. New York: Grune & Stratton, 1952. Pp. 328. (Vol. 1, Sec. 1.)
- LINDNER, R. *Prescription for Rebellion*. New York: Rinehart, 1952. Pp. 305.

Founded by G. Stanley Hall in 1891, reorganized by Carl Murchison in 1925

THE PEDAGOGICAL SEMINARY AND

JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior,
and Comparative Psychology

EDITED BY
CARL MURCHISON

JOHN E. ANDERSON
University of Minnesota

CHARLOTTE BÜHLER
University of
Southern California

CYRIL BURT
University of London

LEONARD CARMICHAEL
Smithsonian Institution

SANTE DE SANCTIS
R. Università di Roma

ARNOLD GESELL
Yale University

WILLIAM HEALY
Judge Baker Foundation,
Boston

Arthur T. Jersild
Teachers College,
Columbia University

BUFORD JOHNSON
The Johns Hopkins
University

HAROLD E. JONES
University of California

TRUMAN L. KELLEY
Harvard University

K. S. LASHLEY
Harvard University

HENRI PIÉRON
Université de Paris

Robert R. Sears
Harvard University

GEORGE D. STODDARD
University of Illinois

CALVIN P. STONE
Stanford University

LEWIS M. TERMAN
Stanford University

GODFREY THOMSON
University of Edinburgh

C. J. WARDEN
Columbia University

JOHN B. WATSON
New York City

If this space should be unstamped, this is the regular library edition.
But if this space is stamped with a designating title, this is a special

edition, sold under the restrictions of a bilateral contract, and may not
be resold for a period of five years from date of publication.

VOLUME 83

1953

Copyright, 1953, by The Journal Press
Published quarterly by The Journal Press
Provincetown, Massachusetts, U. S. A.

AUTOR INDEX

Ades, Harlow W.	63
Altus, Grace T.	241
Bousfield, W. W.	79
Corsini, Raymond J.	249
Ehrenberg, A. S. C.	185
Ellington, R. J.	19
Fassett, Katherine K.	249
Grant, Marguerite Q.	31
Guze, Henry	293
Haggerty, Arthur D.	303
Handlon, Joseph H., Jr.	3
Ives, Virginia	31
Jackson, Joseph	159
Knobloch, Hilda	137
Levy, Nissim	3
Nagy, Maria H.	199, 217, 227
Pasamanick, Benjamin	137
Patton, R. A.	89
Pilgrim, Francis J.	89
Pillsbury, Walter B.	309
Ranzoni, Jane H.	31
Riopelle, Arthur J.	63
Ross, Sherman	179
Schlaegel, T. F., Jr.	265
Seward, 'John P.	3
Springer, Doris	121
Zakolski, F. C.	279
Zeligs, Rose	171

TABLE OF CONTENTS

Choice behavior in satiated rats as a function of drive during training . . .	3
BY JOHN P. SEWARD, NISSIM LEVY, AND JOSEPH H. HANDLON, JR.	
Response to physiological stress in normal and behavior problem children . . .	19
BY R. J. ELLINGSON	
The "neurotic" Rorschachs of normal adolescents	31
BY VIRGINIA IVES, MARGUERITE Q. GRANT, AND JANE H. RANZONI	
Visual discrimination performance in rhesus monkeys following extirpation of prestriate and temporal cortex	63
BY ARTHUR J. RIOPELLE AND HARLOW W. ADES	
The assumption of motor primacy and its significance for behavioral de- velopment	79
BY W. A. BOUSFIELD	
Food consumption and growth of the rat as a measure of motivational stress . . .	89
BY FRANCIS J. PILGRIM AND R. A. PATTON	
National-racial preferences of fifth-grade children in Hawaii	121
BY DORIS SPRINGER	
Further observations on the behavioral development of Negro children . . .	137
BY HILDA KNOBLOCH AND BENJAMIN PASAMANICK	
The effect of classroom organization and guidance practice upon the person- ality adjustment and academic growth of students	159
BY JOSEPH JACKSON	
SHORT ARTICLES AND NOTES	171
Children's concepts and stereotypes of Turk, Portuguese, Roumanian, Arab, Chinese, French-Canadian, Mulatto, South American, Hawaiian, and Australian (Rose Zeligs)	171
A note on dominance tests (Sherman Ross)	179
BOOKS	183
CRITICAL REVIEWS OF RECENT BOOKS	185
Cattell, R. B. <i>Factor Analysis: An Introduction and Manual for the Psy- chologist and Social Scientist.</i> (Reviewed by A. S. C. Ehrenberg)	185
BOOKS RECENTLY RECEIVED	195
Children's conceptions of some bodily functions	199
BY MARIA H. NAGY	
Children's birth theories	217
BY MARIA H. NAGY	
The representation of "germs" by children	227
BY MARIA H. NAGY	

W.I.S.C. patterns of a selective sample of bilingual school children	241
BY GRACE T. ALTUS	
Intelligence and aging	249
BY RAYMOND J. CORSINI AND KATHERINE K. FASSETT	
The dominant method of imagery in blind as compared to sighted adolescents	265
BY T. F. SCHLAEGEL, JR.	
Studies in delinquency: III. An individual test	279
BY F. C. ZAKOLSKI	
Compensatory behavior as completion of the sensory-motor cycle: An inter- pretative hypothesis	293
BY HENRY GUZE	
SHORT ARTICLES AND NOTES	303
The intellectual functioning of a post-institutional group (Arthur D. Haggerty)	303
BOOKS	307
CRITICAL REVIEWS OF RECENT BOOKS	309
Humphrey, G. <i>Thinking</i> . (Reviewed by Walter B. Pillsbury)	309
BOOKS RECENTLY RECEIVED	313

CHOICE BEHAVIOR IN SATIATED RATS AS A FUNCTION OF DRIVE DURING TRAINING*¹

Departments of Psychology, University of California and University of Missouri

JOHN P. SEWARD, NISSIM LEVY, AND JOSEPH H. HANDLON, JR.

A. PROBLEM

This paper reports the second of two closely related experiments on the nature of reinforcement. Let us define reinforcement, following Meehl (8), as the presentation of a stimulus change which, occurring after a response, increases the strength of that response. So as not to beg the question of our research, however, we shall define strength of response in the way one of us has elsewhere suggested (10); simply as the probability that the response will occur in a given situation. This definition also leaves open to empirical determination the kinds of stimulus change that will produce the increased strength. Two kinds of change have been emphasized (3, p. 82): (a) drive reduction, or stimuli associated therewith; and (b) drive increase, as in conditioning the galvanic skin response by means of shock.² For our present purpose we shall confine ourselves to that class of reinforcers associated directly or indirectly with drive reduction.

Leaving aside the question of the nature of reinforcers, we find two major theories of how they operate. Followers of Hull, on the one hand, contend that reinforcement produces an increment of habit strength, or a tendency for a stimulus pattern, under adequate motivation, to arouse a response. Followers of Tolman, on the other hand, hold that reinforcers influence the occurrence of a response by setting up an expectation of its consequences. If the organism expects a response to lead to a positively demanded reinforcer it will be more likely to make that response; if the expected reinforcer carries negative valence the organism will probably make an avoidance reaction instead.

Attempts to verify Tolman's cognition theory have led to the so-called latent learning experiments. In these studies the motivating conditions presumably necessary for reinforcement in the Hullian sense are minimized during training, then introduced in one or more test trials. Thus, hungry rats

*Received in the Editorial Office on May 24, 1951.

¹Aided by a grant from the Faculty Research Committee, University of California.

²It was largely a refusal to subsume the second class of reinforcer under the first that led Mowrer (9) to his two-category theory of learning.

may be permitted to explore an empty maze or be given a series of unrewarded trials, then tested after finding food in the endbox. Or rats satiated for a particular incentive may be run in a Y or T maze with that incentive in one arm, then tested under the appropriate drive. An abrupt shift in the direction of adaptive behavior on test trials is taken to indicate cognition. For an excellent review of such experiments the reader is referred to Thistlethwaite (14); their results may be considered to specify the conditions under which cognitions are or are not formed.

A direct test of Hull's concept of reinforcement calls for a different design. Instead of bringing drive and incentive together *after* training, as in the latent learning experiment, we must reverse this procedure. Animals are first trained to make a response leading to need satisfaction; they are then tested with the need removed. If the reinforcement has built habit strength, this should combine with other, residual drives to produce a somewhat heightened reaction potential (3). But if the reinforcement has merely formed a cognition as to the means of obtaining a reward, then in the absence of a specific demand for that reward we may expect no increased tendency to make the response.

Most of the experiments using this method have found some enhancement of response strength even when the original drive was reduced to zero (for references, see 15). An exception was the study by Koch and Daniel (7), involving a bar-pressing response with food reward. As pointed out by Zeaman and House (15), the extreme satiation procedures employed may have surfeited the animals and thus interfered with any activity. This argument would not apply to maze errors, however, since the negative drive state of a "stuffed" rat should affect both responses at a choice point equally.³ Yet in one of the two maze experiments reported by Siegel (12), errors on satiated trials remained at the chance level throughout. The conditions necessary for the activation of habits in satiated animals are not yet clear.

In this study and the earlier one (11) we have attempted to reduce this problem to a set of simple, unambiguous operations. The first step is to measure the relative strengths of two alternative responses in rats at a choice point. The next step is to force the two responses an equal number of times, always reinforcing one of them but not the other. Finally the two responses are again tested in the absence of the drive rewarded during training, to see if differential reinforcement has caused a shift in habit strength.

Just as in the study of latent learning, there are two ways of manipulating reinforcement. One is to hold drive constant at some positive level and intro-

³Unless one assumes that extreme satiation produces a specific aversion to food.

duce or withhold an appropriate satisfier. The other is to make the satisfier available on all occasions but to raise or lower the drive level. In our first study (11) we adopted the former method. Thirsty rats of the experimental group consistently found water on one side of a choice point but not on the other. When later satiated and tested they chose the reward side significantly more often than before, while a control group showed no change.

In the present experiment we made use of the second method of controlling reinforcement. Rats of the experimental group now found water on both sides of the choice point, but whenever they found it on one side they were already satiated, whereas on the other side they were invariably thirsty. The empirical question at issue was whether, when satiated and given a free choice, these rats would show an enhanced preference for the "thirsty" side.

Fundamentally, the theoretical problem underlying both experiments is the same: to determine the effect of reinforcement on habit strength apart from its effect on cognition. Compared with each other, however, the results of the two methods should throw light on a second problem; namely, the relative variation of reinforcement as a function of drive and of incentive. It is conceivable that, with drive active, the presence or absence of a satisfier may be highly significant for habit strength; whereas, if the satisfier is available, the presence or absence of drive may make little difference. The experiment here reported thus raises a specific problem within the framework of the more general one. In the context of Hull's system this further problem may be formulated as follows: *to determine the relative effectiveness of a secondary reinforcer (water) with and without the primary drive (thirst)*. The two problems are partially independent. A positive result in the present experiment would necessarily apply to both the more general and the more specific problem. On the other hand, a finding of no difference between the two conditions of reinforcement would not imply the ineffectiveness of reinforcement as a whole.

B. METHOD

1. Design

The experimental design consisted essentially of two free-choice tests separated by a training period. The tests were run under satiation for food and water. After the first test the rats were divided into four groups, two experimental and two control, matched for position preference. They were then given 12 days of training, on half of which they were thirsty and on the other half satiated for water. Water was available on both sides of the choice point. On thirsty days the experimental groups were permitted to

practice only the initially weaker of the two responses, on satiated days only the stronger one. The control groups, on the other hand, divided their practice on both responses equally between satiated and thirsty days.

Four subgroups were required in order to keep constant all other conditions of training. For example, one experimental subgroup might be trained two days thirsty to the nonpreferred side, two days satiated to the preferred side, and so on. A control subgroup would follow the same drive schedule but every second day it would be practiced on the opposite response. Since rate of alternation of response might influence performance on the final preference test, a second experimental subgroup was required in which both drive state and response were shifted daily. This in turn required a second control subgroup with daily alternation of thirst and satiation but with response alternation every two days. Furthermore, results on the first sample (*A*) suggested that the schedule on the last days of training might affect performance on the final preference test. A second sample (*B*) was therefore run with the order of training completely reversed. Finally, in an attempt to reduce inter-group variance, a third sample (*C*) was run with a more irregular pattern of alternations. The entire schedule is shown in Table 1.

2. Subjects

The subjects were male albino and hooded rats from the U. C. L. A. colony, from three to five months old at the beginning of the experiment. Data are reported on 80 rats, 24 in Sample *A*, 24 in Sample *B*, and 32 in Sample *C*.⁴ A total of 105 rats were started, but five were dropped in the preliminary stage due to illness or poor adaptation and 20 more were discarded in the process of matching.

3. Apparatus

For a detailed description with a drawing of the choice box see (11). It consisted essentially of three parallel alleys 24 inches long, by 4 inches wide, by 5½ inches high. A sliding metal door divided the central alley into a starting box and a choice area; the latter gave access to the outer alleys through two transparent, one-way doors held open by the experimenter and dropped behind the animal to prevent retracing. On forced trials one of these doors was locked shut. At the far end of each outer alley was a metal cup for the incentive. The entire box was covered by a one-way-vision roof. A strip of movable oilcloth on the floor prevented tracking.

⁴Samples *A* and *B* were run by J. H. H.; Sample *C*, by N. L.

TABLE 1
TRAINING SCHEDULE*

		Days											
Sample	Group	1	2	3	4	5	6	7	8	9	10	11	12
<u>A</u>	E1	S P	S P	T N	T N	S P	S P	T N	T N	S P	S P	T N	T N
	C1	S P	S N	T P	T N	S P	S N	T P	T N	S P	S N	T P	T N
	E2	S P	T N	S P	T N	S P	T N	S P	T N	S P	T N	S P	T N
	C2	S P	T P	S N	T N	S P	T P	S N	T N	S P	T P	S N	T N
<u>B</u>	E1	T N	T N	S P	S P	T N	T N	S P	S P	T N	T N	S P	S P
	C1	T N	T P	S N	S P	T N	T P	S N	S P	T N	T P	S N	S P
	E2	T N	S P	T N	S P	T N	S P	T N	S P	T N	S P	T N	S P
	C2	T N	S N	T P	S P	T N	S N	T P	S P	T N	S N	T P	S P
<u>C</u>	E1	T N	S P	S P	T N	T N	S P	S P	T N	T N	S P	S P	T N
	C1	S N	T P	S P	T N	T N	S P	T P	S N	S N	T P	S P	T N
	E2	S P	T N	T N	S P	S P	T N	T N	S P	S P	T N	T N	S P
	C2	T P	S N	T N	S P	S P	T N	S N	T P	T P	S N	T N	S P

*E1, E2—experimental subgroups; C1, C2—control subgroups. The top line for each group gives the drive sequence: S—satiated, T—thirsty. The bottom line gives the response sequence: P—preferred side, N—nonpreferred side.

4. Procedure

a. *Adaptation.* In order to induce more active movement during later satiated trials, adaptation to the choice box was carried out under hunger-food motivation.⁵ All food was removed from the cages and mash was provided for one hour (in some cases 45 minutes) daily. After four to six days on this diet, with daily handling, the rats were allowed four or five days of group exploration, 10 minutes a day; they were put in the choice box in groups of three to five with incentive cups empty and all doors open. This was followed by 10 to 13 days of individual adaptation, until most animals ran readily. Each rat was given two forced trials a day, one to each side, 15 to 30 minutes apart, close to the usual feeding time. Mash and large pellets were used as incentives. Half the rats started to the right and half to the left; for each animal the order of right and left trials was reversed daily.

b. *Free-choice Test 1 (satiated).* Before testing for position preference the rats were satiated for food and water. This was done by leaving pellets and water bottles in the cages except for the period from four hours to one hour before testing. At the close of that period cups of wet mash and water were added until the trials were over. In later groups the deprivation period was increased to the interval between five hours and one-half hour before the test began.

Four free trials a day were given on two successive days. Four to eight rats at a time were given their trials in rotation, providing an interval of 15 to 30 minutes between trials. Although animals were allowed about 15 seconds in the chosen alley, where they found both water and pellets, no rat was observed to eat or drink in the choice box during a satiated preference test.

c. *Training.* On the basis of proportion of choices to one side in the preceding test, the rats were sorted into sets of four, each set consisting of animals with approximately equal bias. The four rats of each set were then assigned at random to one of the four groups, two experimental and two control, mentioned above. Litter mates were matched when possible.

The training schedule has already been presented in Table 1. Each animal was given four forced trials a day, all to the same side, 15 to 30 minutes apart. On thirsty days he was allowed to drink for about 15 seconds from

⁵It may be noted that this procedure is based on the hypothesis under test; namely, that reinforcement builds habit strength for possible use in the service of other drives. Although in one sense it begs the question, in another it does not, since it leaves the hypothesis open to verification by the experiment proper.

the incentive cup; on satiated days he was left in the outer alley the same length of time.

Thirst was controlled as follows: On thirsty days water was removed 21 hours before trials were run, while pellets were continuously available; water and wet mash were supplied for one hour immediately after trials were over. On satiated days both water and pellets were kept available. Five hours before training time water was removed for $4\frac{1}{2}$ hours; cups of water were then added to the customary bottles and the dry food was removed until all trials had been run.⁶

d. Free-choice Test 2 (satiated). On the next two days after training, all rats were again tested for position preference under the same conditions as the first free-choice test.

e. Free-choice Test 3 (thirsty). This test was not included in the original design. Since Test 2, however, approximated the drive conditions of training on satiated days, it was thought desirable to hold a test under the same drive conditions as on thirsty days. On the next two days after the second test, therefore, all rats were deprived of water for 21 hours and given four free trials a day in the same manner as on the previous tests.

Throughout the experiment all trials were timed from the raising of the starting door until the rat passed the midline of the chosen alley. If a trial took longer than three minutes the rat was sometimes urged gently forward; this never occurred at the choice point on free-choice trials.

C. RESULTS

1. Frequencies of Choice

Table 2 gives a bird's-eye view of the main results for all samples combined.

TABLE 2
MEAN NUMBER OF CHOICES OF INITIALLY PREFERRED SIDE ON TESTS 1 AND 2 (SATIATED) AND 3 (THIRSTY) FOR EXPERIMENTAL AND CONTROL GROUPS, BASED ON ALL SAMPLES: DIFFERENCES BETWEEN TESTS

		DIFFERENCES BETWEEN TESTS					
Group	N	Tests			Differences		
		1	2	3	1-2	1-3	2-3
E	40	5.85	3.18	1.90	2.67	3.95	1.28
C	40	5.85	5.00	4.54*	0.85	1.36**	0.51**
E-C		0.00	-1.82	-2.64	1.82	2.59	0.77

*N = 39; one rat sick.

**These differences do not correspond with means in left half of table since sick rat's scores on earlier tests also omitted here.

⁶A less extreme satiation technique, with water removed for only three hours and pellets always present, was corrected when some rats in Sample A were seen to drink in the choice box on satiated days. The revised procedure eliminated this behavior.

In the left half of the table appear the mean frequencies of choice of the initially preferred side on all three free-choice tests by the experimental and control groups. In the right half are shown the differences between pairs of tests. The bottom row gives the differences between Groups *E* and *C*. It will be observed that the two groups were equal at the start and that both fell off somewhat from test to test. But while Group *C* continued to average more than half of its choices to the preferred side, Group *E* reversed its preference on Test 2 and accentuated that reversal on Test 3. It is noteworthy that the groups diverged more markedly between the first and second satiated tests than between the second satiated and final thirsty test.

Before finding the significance of these differences it was necessary to evaluate the other main sources of non-random variation in the experimental design. These resulted from the use of three samples of rats, 12 different training orders (see Table 1), and sets of animals matched for initial preference. The last factor was ignored when it was found that the correlation between preference scores on Test 1 and *shifts* of preference on Test 2 was negligible. The other two factors, however, were not so readily dismissed.

Tables 3 and 4 provide a breakdown of the data in Table 2 by samples

TABLE 3
MEAN NUMBER OF CHOICES OF INITIALLY PREFERRED SIDE ON TESTS: DATA FROM LEFT HALF OF TABLE 2 BROKEN DOWN BY SAMPLES AND TRAINING ORDERS*

Group	Sample A			Sample B			Sample C		
	1	2	3	1	2	3	1	2	3
E1	6.17	5.33	3.67	6.17	2.50	0.17	5.63	3.00	3.38
E2	6.00	2.67	0.33	5.83	2.83	0.17	5.50	2.88	2.88
C1	6.33	5.00	2.83	6.00	6.83	5.83	5.63	5.00	5.13
C2	5.83	2.83	2.67	5.83	5.67	6.20**	5.63	4.75	4.63

*Means based on six cases in Samples *A* and *B*, eight in Sample *C*.
**N = 5, since one rat sick on Test 3.

TABLE 4
DIFFERENCES BETWEEN TESTS IN MEAN NUMBER OF CHOICES OF INITIALLY PREFERRED SIDE: DATA FROM RIGHT HALF OF TABLE 2 BROKEN DOWN BY SAMPLES AND TRAINING ORDERS

Group	Sample A			Sample B			Sample C		
	1-2	1-3	2-3	1-2	1-3	2-3	1-2	1-3	2-3
E1	0.84	2.50	1.66	3.67	6.00	2.33	2.63	2.25	-0.38
E2	3.33	5.67	2.34	3.00	5.66	2.66	2.62	2.62	0.00
C1	1.33	3.50	2.17	-0.83	0.17	1.00	0.63	0.50	-0.13
C2	3.00	3.16	0.16	0.16	0.00*	0.00*	0.88	1.00	0.12

*These differences do not correspond with means in Table 3 since sick rat's scores on earlier tests also omitted here.

and training orders. Although the trend already noted is clearly discernible in Samples *B* and *C*, it is obscured in Sample *A* and there is considerable variation among the subgroups.

The difference scores between tests, as the best indices of the effects of training, were selected for analysis of variance. Greatest interest attached to the differences between Tests 1 and 2 in the experimental and control groups. Bartlett's test for homogeneity of variance (2) applied to the 12 subgroups yielded a χ^2 of 13.06, significant at the .30 level. The results of the analysis appear in Table 5.

TABLE 5
ANALYSIS OF VARIANCE OF DIFFERENCES BETWEEN TEST 1 AND 2 IN NUMBER OF CHOICES
OF INITIALLY PREFERRED SIDE

Source	$\Sigma\chi^2$	df	s^2	F	p
Experimental-Control	66.6	1	66.6	27.75	<.001
Samples	5.0	2	2.5	1.04	—
Groups X Samples					
Interaction	42.2	2	21.1	8.79	<.001
Order within Samples	31.7	6	5.3	2.21	.05
Within subgroups	165.0	68	2.4		

Using the within-cells variance as an error estimate we find two *F*s significant at better than the .001 level; the variance between Groups *E* and *C* and the interaction of this factor with the variance between samples. The residual variance between training orders is barely significant at the .05 level. If we test the *E-C* variance against its significant interaction we obtain $F=3.2$, as compared with the *F* of 18.51 required for significance at the .05 level. It follows that the differential treatment given Group *E* produced a real effect on its satiated position preferences, but that this effect cannot be generalized to all training orders.⁷

Although the divergence of the two groups seemed to be still further accentuated under drive, a complete analysis of variance of the differences between Tests 1 and 3 was precluded by the fact that Bartlett's test gave a χ^2 significant at the .01-.02 level. Inspection of the subgroup variances showed that the heterogeneity, caused by the relatively wide dispersion of the control scores, was mainly confined to Samples *A* and *B*. Treatment of the Sample *C* data alone by analysis of variance yielded an *F* for the *E-C* variable equal to 10.40, which with one and 28 degrees of freedom is significant at the .01 level. The differences between Tests 2 and 3, however, failed to distinguish significantly between Groups *E* and *C*. Since Bartlett's χ^2 test remained above

⁷This inference involves the assumption that training order was the variable responsible for the interaction with samples.

the .10 level we ran an analysis of variance for the entire population, but only the F for samples reached the .05 level of significance. We may infer that the effect of differential reinforcement in this experiment was largely independent of whether it was tested under drive or satiation.

2. Running Times

Choice frequency reflects only a shift in the *relative* strengths of two responses. To throw light on the nature of the underlying changes we therefore turned to another index of response strength. Running time was defined as the time from raising the starting door until the animal passed the midline of the chosen alley. Because the distribution of this measure was highly skewed, and because not all trials were completed within the arbitrary time limit, we used the median to represent central tendencies.

Our first question was whether running times would show any trends in the course of training. To answer it we divided the 12 days of training into three sections of four days each. (As shown in Table 1, all training orders could be divided into four-day cycles, each containing an equal number of all drive-response combinations used.) We then pooled all trials within a section and found the median. This was first done separately for each drive-response combination (SP , SN , TP , TN), but since the control subgroups showed no differences between preferred and nonpreferred responses under the same drive state, this distinction was abandoned. Thirsty and satiated times, however, were clearly demarcated and were therefore kept separate. By the same logic the data of Samples A and B were pooled while Sample C was treated as a unit. Medians were also computed for the three tests by pooling all eight trials of all animals.

Table 6 presents the median running times on Tests 1 and 2 and on satiated days during each section of training, while Table 7 does the same for the the thirsty days of training and Test 3. On the whole the two tables reveal opposite trends shared alike by experimental and control groups. While sat-

TABLE 6
MEDIAN RUNNING TIMES ON SATIATED TEST AND SATIATED TRAINING DAYS

Sample	Group	N	Test 1	Training Days			Test 2
				1-4	5-8	9-12	
A & B	E	192	24.4	19.0	23.3	27.9*	27.7
	C	192	22.4	21.5	20.1	25.9	28.2
C	E	128	17.9**	39.5	72.4	63.5	57.5**
	C	128	31.0**	36.8	59.5	56.2	89.0**

*N = 188, since one rat sick one day.

**N = 64, since times on Test 2 recorded for only half of the animals in Sample C.

TABLE 7
MEDIAN RUNNING TIMES ON THIRSTY TRAINING AND THIRSTY TEST DAYS

Sample	Group	N	Training Days			Test 3*
			1-4	5-8	9-12	
A & B	E	192	5.8	2.9	2.8**	3.8
	C	192	7.7	3.4	3.4	4.2
C	E	128	10.0	5.2	3.4	—
	C	128	7.5	4.9	3.8	—

*On drive test for Samples A and B, N = 184 and 183 for Groups E and C, respectively, due largely to sickness of one C rat and omission of his E partner. Times omitted for Sample C since recorded for only half of the animals.

**N = 184; one rat sick two days.

iated times tended to become longer as the experiment proceeded, thirsty times were reduced.

Testing the statistical significance of these trends presented some difficulty, since they were derived from distributions that violated assumptions underlying the usual tests. The method finally hit upon was as follows: For each rat we determined the median running time for a given test or section of training. We then tested the null hypothesis that between any two phases of the experiment as many rats should increase their time as decreased it. For the purpose of this test all groups and samples were pooled. The results appear in Table 8. It will be noted that except for the period between the second and third sections of training the consistency of both the upward-satiated and downward-thirsty trends is highly significant.

Further evidence of the opposite effects of satiation and thirst on response strength was obtained by a comparison of running times from trial to trial within the same day. For this purpose Groups E and C were combined. All

TABLE 8
NUMBER OF RATS INCREASING SATIATED AND DECREASING THIRSTY MEDIAN RUNNING TIMES IN SUCCESSIVE PHASES OF EXPERIMENT, WITH LEVELS OF SIGNIFICANCE: ALL GROUPS AND SAMPLES COMBINED

GROUPS AND SAMPLES COMBINED									
Phases		N	Satiated			N	Thirsty		
			Incr.	z^*	p^{**}		Decr.	z	p
Thirds of training:	I-II	80	61	4.59	<.0001	80	70	6.60	<.0001
	I-III	80	61	4.59	<.0001	79	70	6.74	<.0001
	II-III	80	45.5†	1.12	.30	79	51	2.47	.02
	1-2	64	45	3.13	.002				
Tests									

$$*z = \frac{|X - np| - 0.5}{\sqrt{npq}}$$

**Two-tailed test.

†A rat that had the same median in both phases was scored 0.5.

first trials of all rats in the 12 days of training were pooled and the median obtained; the same was done for second, third, and fourth trials. This procedure was carried out twice, once for satiated days and once for thirsty.

TABLE 9
MEDIAN RUNNING TIMES ON SUCCESSIVE TRIALS WITHIN A DAY DURING TRAINING:
EXPERIMENTAL AND CONTROL GROUPS COMBINED

Sample	N	Satiated Trials				N	Thirsty Trials			
		1	2	3	4		1	2	3	4
A & B	287	12.9	21.9	29.2	33.9	286	6.3	3.9	3.1	3.0
C	192	16.7	43.5	64.9	95.3	192	6.4	5.2	4.4	3.9

To judge by the medians in Table 9, when satiated the rats tended to run more slowly on successive trials; when thirsty they tended to accelerate. To test the significance of these trends we used the method already described; i.e., we determined how many of the 80 rats shifted in the specified direction from one trial to each subsequent trial in turn. Of the 12 frequencies computed, 11, ranging from 60.5 to 77, exceeded chance at $p < .0001$. The one exception—only 52 thirsty rats lowered their median times between the third and fourth trials—barely reached the .01 level.

D. DISCUSSION

We started with the question: Does a reinforcer increase habit strength? To answer this question we first provided differential reinforcement for two responses, then tested its effect with the relevant drive removed, thus presumably nullifying any influence from sheer cognition. "Differential reinforcement" in this experiment consisted of "rewarding" one response in the presence of the relevant drive, the other in its absence. This technique raised a second question: Is a secondary reinforcer more effective in combination with the relevant drive than without it?

Our results indicate a positive answer to both questions. After finding water on one side of a choice point when thirsty and on the other side when satiated, the experimental rats showed a significant shift of preference to the thirsty side, even under satiation, as compared with the controls.

As it stands, however, this finding may not be entirely satisfactory, even to a reinforcement theorist. In the first place, he may point out that Hull (3, 5) does not explicitly include drive in the expression for habit strength and that Strassburger (13) found no relation between resistance to extinction and drive strength during learning. It should be noted, however, that we did not measure the effect of varying supraliminal drive intensity but compared

one level of drive with satiation. Our data do not conflict, therefore, either with the theory or with the evidence cited; they simply indicate that statements about habit strength as a function of drive intensity cannot be extrapolated to zero drive.

In the second place, even granting the greater effectiveness of incentive-plus-drive over incentive alone, a critic may question the logic of our prediction. Test 2, he may point out, was run under essentially the same drive condition as the satiated training trials; there is experimental evidence (1, 6), moreover, that drive intensities can furnish clues to discriminative learning. Why, then, should we not expect a satiated rat to choose the side on which he had been reinforced when satiated?

An answer is to be found in the running times. These data suggest that in spite of the similarity of external conditions the animals were learning two different things when thirsty and when satiated. On drive days they ran more rapidly from trial to trial while on satiation days they slowed down. Some such concept as "negative adaptation" or "reactive inhibition" (3) might account for the latter finding, but it would not account for the cumulative retardation from day to day. For this purpose we might suppose a gradual approach to maximal satiation, but such an assumption is purely *ad hoc* and receives no support from the parallel *acceleration* on intervening days under drive. A more obvious interpretation is that the rats were learning to run when thirsty and not to run when satiated.

The running time data enable us to attempt a rudimentary account of our finding in terms of Hull's theory. They provide a basis for the statement that the experimental conditions on drive days and satiation days were such as to produce habit strength (sH_R) and conditioned inhibition (sI_R), respectively. That is, training on thirsty days strengthened the habit of running in response to a stimulus pattern coming essentially from the choice box and from the internal state of thirst. Satiation training, on the other hand, conditioned an inhibitory process to the stimulus pattern produced by the choice box and the internal state of satiation. A discriminative reaction was thus acquired to the presence or absence of the thirst stimulus.

So far the description applies equally to Groups *E* and *G*. But the experimental rats, unlike their controls, were given the opportunity to develop a second discrimination. It will be recalled that during training they responded to the thirst stimulus in conjunction with external cues predominantly from the nonpreferred side, since the door was open on that side only. The satiation stimulus, on the other hand, was combined chiefly with cues from the opposite side. Conditions were thus provided for nonpreferred cues to acquire

some degree of control over the running habit and for preferred cues to become primarily conditioned to the inhibitory response. Along with these developments we must recognize the likelihood that each learning process gave rise to a family of generalization gradients extending along the various stimulus and response dimensions involved.

We now face the question of how these acquired tendencies would be expected to resolve themselves in an *E* rat at the choice point during his second satiated test. To account for our results we must show that effective reaction potential ($s\bar{E}_R$) should be greater to the nonpreferred side than to the preferred. A rigorous proof would be too complicated. Since in this situation $s\bar{E}_R$ would seem to vary largely as a function of sH_R and sI_R , we shall simply try to demonstrate the above relationship for the algebraic sums of these factors.

The following notations will simplify our task: Let us denote habit strength to either side before training by H_o (since we are concerned with *changes* in strength, initial bias may be ignored). Let us also denote stimuli on the nonpreferred and preferred sides by S_n and S_p , respectively; responses to those stimuli by R_n and R_p ; and internal stimuli due to thirst and satiation by S_{D_i} and S_{D_o} .

On the foregoing assumptions, training under thirst will result in the following additions to H_o : (a) by direct reinforcement, $s_n + s_{D_i}H_{R_n}$; by stimulus-response generalization, (b) $s_n + s_{D_o}H_{R_n}$ and (c) $s_n + s_{D_o}H_{R_p}$. But by virtue of the generalization gradient,

$$s_n + s_{D_o}H_{R_n} > s_p + s_{D_o}H_{R_p}. \quad (1)$$

Training under satiation will result in the following *subtractions* from H_o : (a) by direct conditioning, $s_p + s_{D_o}I_{R_p}$; (b) by stimulus-response generalization, $s_n + s_{D_o}I_{R_n}$. But, again by virtue of the gradient,

$$s_p + s_{D_o}I_{R_p} > s_n + s_{D_o}I_{R_n}. \quad (2)$$

On Test 2 these factors will enter the competing reaction potentials at the choice point as follows:

$$s_n E_{R_n} = f (H_o + s_n + s_{D_i}H_{R_n} - s_n + s_{D_o}I_{R_n}), \text{ and}$$

$$s_p E_{R_p} = f (H_o + s_p + s_{D_o}H_{R_p} - s_p + s_{D_o}I_{R_p})$$

But from inequalities (1) and (2) it follows that

$$s_n E_{R_n} > s_p E_{R_p},$$

from which the choice behavior of our experimental rats may be predicted.

In the same approximate manner it may be shown: (a) that Group *E* rats should also tend to choose their initially nonpreferred side on Test 3 under drive, and (b) that in neither case would such a change be predicted for Group *C*. It may be pointed out that the foregoing account makes no mention of afferent stimulus interaction (4, 5) induced by the shift from forced training trials to free test trials and from water satiation during training to food *and* water satiation on Test 2. There is no reason to believe, however, that its inclusion would alter the predicted outcome. We conclude, therefore, that at the qualitative level of analysis here attempted our findings can be readily deduced from Hullian postulates.

E. SUMMARY

To test the effect of reinforcement on habit strength aside from cognition, 80 satiated rats were given an initial position preference test in a choice box with food and water on both sides. Two matched groups of 40 each were then given four forced trials a day for 12 days, on six of which they were thirsty and on the other six satiated for water. For Group *E*, all trials on thirsty days were to the nonpreferred side; on satiated days, to the preferred side. For Group *C* all trials on half the thirsty days and half the satiated days were to the nonpreferred side; on the other half, to the preferred. Water was present on both sides during training. All rats were then tested for position preference as before, once satiated for food and water, once thirsty.

The following results were obtained:

1. As shown by the mean number of choices of the side originally preferred, Group *E* shifted its preference to the opposite side on both subsequent tests; Group *C* shifted relatively little. The difference between the two groups in amount of shift from Test 1 to 2 was significant at $p < .001$; that between two subgroups from Test 1 to 3, at $p < .01$. From Test 2 to 3 the groups did not differ significantly in amount of shift.

2. Median running times under satiation for both Groups *E* and *C* increased from trial to trial on a single day and from early to late in training. Median running times under thirst decreased during the same intervals. The consistency of these trends was well beyond chance expectancy.

A Hullian interpretation in terms of the conditioning and generalization of habit strength and inhibition was suggested. Our findings point to two conclusions, the generality of which remains to be established: (a) Reinforcement increases habit strength apart from cognitions. (b) A secondary reinforcer is more effective in the presence of a relevant drive than in its absence.

REFERENCES

1. BLOOMBERG, R., & WEBB, W. B. Various degrees within a single drive as cues for spatial response learning in the white rat. *J. Exp. Psychol.*, 1949, **39**, 628-636.
2. EDWARDS, A. L. *Experimental Design in Psychological Research*. New York: Rinehart, 1950. Pp. 446.
3. HULL, C. L. *Principles of Behavior*. New York: Appleton-Century, 1943. Pp. 422.
4. ———. The discrimination of stimulus configurations and the hypothesis of afferent neural interaction. *Psychol. Rev.*, 1945, **52**, 133-142.
5. ———. Behavior postulates and corollaries—1949. *Psychol. Rev.*, 1950, **57**, 173-180.
6. JENKINS, J. J., & HANRATTY, J. A. Drive intensity discrimination in the albino rat. *J. Comp. & Physiol. Psychol.*, 1949, **42**, 228-232.
7. KOCH, S., & DANIEL, W. J. The effect of satiation on the behavior mediated by a habit of maximum strength. *J. Exp. Psychol.*, 1945, **35**, 167-187.
8. MEEHL, P. E. On the circularity of the law of effect. *Psychol. Bull.*, 1950, **47**, 52-75.
9. MOWRER, O. H. On the dual nature of learning—a re-interpretation of "conditioning" and "problem-solving." *Harv. Educ. Rev.*, 1947, **15**, 102-148.
10. SEWARD, J. P. Secondary reinforcement as tertiary motivation: a revision of Hull's revision. *Psychol. Rev.*, 1950, **57**, 362-374.
11. SEWARD, J. P., & HANDLON, J. H. The effect of satiation on the use of habit. *J. Genet. Psychol.*, 1952, **81**, 259-272.
12. SIEGEL, P. S. Drive shift, a conceptual and experimental analysis. *J. Comp. Psychol.*, 1943, **35**, 139-148.
13. STRASSBURGER, R. C. Resistance to extinction of a conditioned operant as related to drive level at reinforcement. *J. Exp. Psychol.*, 1950, **40**, 473-487.
14. THISTLETHWAITE, D. L. A critical review of latent learning and related experiments. *Psychol. Bull.*, 1951, **48**, 97-129.
15. ZEAMAN, D., & HOUSE, B. J. Response latency at zero drive after varying numbers of reinforcements. *J. Exp. Psychol.*, 1950, **40**, 570-583.

Department of Psychology
 University of California
 Los Angeles 24, California

Department of Psychology
 University of California
 Berkeley 4, California

RESPONSE TO PHYSIOLOGICAL STRESS IN NORMAL AND BEHAVIOR PROBLEM CHILDREN*

The Mooseheart Laboratory for Child Research

R. J. ELLINGSON¹

A. INTRODUCTION

Children and adults with behavior disorders frequently show deviations from the norm in a variety of physiological processes (2, 3, 6, 7, 12-15). Although environmental factors are undoubtedly of prime importance in the etiology of many behavior disorders there is some reason to believe atypical physiological reactions, especially under stressful conditions, may predispose to the development of abnormal behavior patterns.

The present study has sought to extend this line of investigation by comparing the physiological responses of four groups of children under "basal" and stressful conditions. The four groups were: (a) a normal or control group, (b) behavior problem children, (c) "normal" children with one schizophrenic parent, (d) "normal" children with one epileptic parent. The physiological stress conditions were hyperventilation and the cold pressor test. The physiological measures were the electroencephalogram, electromyogram, galvanic skin response, respiration, and heart rate.

B. SUBJECTS

All of the children in this study are residents of the "Child City" of Mooseheart, Illinois, established and maintained by the Loyal Order of Moose for children of deceased members of the Order. All were within the age range 10 to 14 years. *Group C — Normal control subjects*, consisted of 5 boys and 5 girls, without personal or family history of behavior or neurological disorders. *Group B — Behavior problem children*, was composed of 7 boys and 3 girls representing the most severe behavior problems within the age range of this study. None had physical defects nor family histories of psychiatric or neurological disorders. One pair of male siblings were monozygotic twins. *Group S — Children with one schizophrenic parent*, composed

*Received in the Editorial Office on June 1, 1951.

¹This study is part of a dissertation submitted for the Ph.D. degree in Psychology at Northwestern University under the direction of Prof. Donald B. Lindsley. The work was carried out in the Mooseheart Laboratory for Child Research; facilities were made available by Dr. Martin L. Reymert, Director.

of 6 boys and 6 girls, including three pairs of siblings. Eleven of these children were behaviorally normal; one was a problem child and was omitted from the group in statistical comparisons. *Group E* — *Children with one epileptic parent*, composed of 6 boys and one girl, including two pairs of siblings. Five of the group were behaviorally normal; two were behavior problems and were omitted from the group in statistical comparisons of the groups.

The mean ages and *IQ*s of the groups were as follows: *Group C*, 12 yrs. 0 mos.; *IQ* 102; *Group B*, 12 yrs. 2 mos.; *IQ* 91; *Group S*, 12 yrs. 3 mos.; *IQ* 103; *Group E*, 12 yrs. 5 mos.; *IQ* 112.

C. PROCEDURE

Electrical phenomena were recorded with a four-channel Offner electroencephalograph unit, Model *A*. These consisted of the electroencephalogram (*EEG*) from the left occipital region, scalp to ear leads; the electromyogram (*EMG*) from frontalis muscle, above nasion, with reference lead over zygoma; galvanic skin response (*GSR*)² from leads one inch apart in center of left palm; electrocardiogram (*EKG*) from left hand and ear leads, from which heart rate was derived. Respiration was recorded pneumographically on the same record.

The subject lay supine in a dimly lit, electrically shielded room, under constant observation. Every effort was made to reduce apprehension. Short strips of record were taken at intervals until physiological measures appeared to be at a "basal level," when a 60-second strip of "basal record" was taken. Hyperventilation at about 30 respirations per minute then followed for a period of 90 seconds, with recording continuing until a basal state was reestablished. A minimum interval of 15 minutes was allowed before the cold pressor test was begun. Following the method of Hines and Brown (4) the right hand was immersed to the wrist in water at 4°C. for 60 seconds; recording was continued until a basal state was reestablished.

Each of the recordings during the three conditions, "*basal*" period of 60 seconds duration, *hyperventilation* — from 60 seconds before to recovery, and *cold pressor test* — from 60 seconds before to recovery, was divided into 10-second sections. These were then analyzed for each measure. Heart rate was determined in beats per minute. A *GSR* Index was derived by tracing the *GSR* curve with a map-measurer; the excess over the linear distance for 10 seconds was the *GSR* Index in units of tenths of an inch. This value was

²*GSR* was recorded as a skin potential (Tarchanoff effect), in contrast to the usual resistance change (Fere effect). The two techniques yield similar results (8), a fact confirmed by preliminary investigation to be reported elsewhere.

corrected for differences in amplification in order to make unit values equal from subject to subject. Respiration was measured in cycles per minute, except during hyperventilation; recovery from hyperventilation was measured as the duration from end of deep breathing to the return of the basal respiratory pattern. The *EMG* was analyzed in terms of the average voltage in microvolts. Five *EEG* measures were derived: percent-time alpha, percent-time 3 to 8 per second waves, alpha frequency, alpha voltage, and delta index for waves of 3 per second or slower, after the method of Hoagland *et al.* (5). Finally, a hypersynchrony rating for slow waves during hyperventilation was established on a 3-point scale, with 1 representing "no 3 per second slow waves," 2 representing "a few scattered slow waves," and 3 standing for short runs or continuous high voltage 3 per second waves.

D. RESULTS

1. Resting or Basal State Records

Table 1 shows the resting or "basal state" values for each of the measures. Group *C* had a higher average resting heart rate and respiratory rate than

TABLE 1
MEAN BASAL ACTIVITY

Measure	Group			
	C	S	B	E
Heart Rate, beats per min.	85.9	76.9	78.4	78.9
Respiration, per min.	19.8	18.5	17.4	16.5
Muscle Potential (μ V)	6	9	8	6
Percent-time alpha (EEG)	74.0	60.5	50.5	53.4
Percent-time 3-8/sec. (EEG)	13.7	16.4	15.8	20.8
Delta Index (EEG)	4.1	4.8	3.8	5.3
Alpha Frequency (EEG)	10.1	10.2	9.2	9.7
Alpha Voltage (EEG)	54	42	46	37
N	10	11	10	5

any of the other groups. The control group also showed higher percent-time alpha and alpha voltage, and lower percent time 3-8 activity and delta index than the experimental groups. However, none of these differences in means is statistically significant, except the difference in alpha frequency between Groups *C* and *B* (significant at the 2 per cent level). On the whole it was felt that none of the measures during the resting or basal period, including the *EEG*, differentiate among the four groups of subjects. This points up the necessity for imposing stress to bring out physiological differences.

2. Reaction to Stress Induced by Hyperventilation

a. *Hypersynchrony of the EEG.* Since data on the effect of hyperventilation upon the EEG were available for 19 additional normal subjects of the same age range as Group C, these were added to provide a larger control group for this aspect of the study. Of the 29 subjects now composing Group C only two, or 7 per cent, showed a degree of hypersynchrony which could be given a rating of 3 according to the scale for rating slow waves during hyperventilation. In contrast, experimental Groups S, B, and E produced, respectively, 36, 33, and 40 per cent with slow waves of the same rating on the hypersynchrony scale. Thus it is evident that there was a significantly greater proportion of experimental (Groups S, B, and E) than control (Group C) subjects showing the development of hypersynchronous slow waves within the 90-second period of overbreathing. The percentage of normals giving this hyperventilation response is about the same as that found by Brill and Seidemann (1) among the 10- to 14-year-olds of their normal group.

b. *Other reactions to hyperventilation.* Table 2 shows for each of the subject groups the differences in the various measures prior to and during hyperventilation. In addition it shows the differences before and during hyperventilation for all subjects, irrespective of group, when segregated according to the criterion of whether or not they showed hypersynchronous slow waves as a result of hyperventilation.

None of the EEG measures shown in the table differentiates significantly

TABLE 2
MEAN RESPONSES TO HYPERVENTILATION: ACTIVITY DURING HYPERVENTILATION
MINUS BASAL ACTIVITY

Measure	Group					
	C	S	B	E	Hyper- synch.	No Hyper- synch.
Heart Rate, beats per min.	13.3	21.6	20.0	14.8	19.0	16.0
GSR Index	4.8	5.8	8.2	11.6	10.6	5.9
Muscle Potential (μ V)	15.1	8.5	3.1	27.4	12.1	10.7
Percent-time alpha (EEG)	-4.9	-4.0	-9.1	1.8	-13.4	-3.2
Percent-time 3-8/sec. (EEG)	5.4	6.9	10.5	6.0	14.9	4.1
Delta Index (EEG)	2.6	3.9	5.4	5.2	8.9	1.9
Alpha Frequency	0.17	-0.15	-0.10	0.28	0.02	0.14
Alpha Voltage (EEG)	2.4	4.1	-0.6	10.0	5.9	2.8
Respiratory Recovery Time, sec.	152	138	197	216	183	154
N	10	11	9	5	9-13	20-24

between the various subject groups; however, percent-time alpha, percent-time 3-8 per second waves, and delta index measures differentiate significantly (1 per cent level or better) between the hypersynchronous and non-hypersynchronous categories, a fact which confirms quantitatively the hypersynchrony ratings mentioned above.

Among the other measures shown in the table, only heart rate can be shown to have differentiating value. The heart rate changes as a result of hyperventilation differ significantly between Groups *C* and *S* (2 per cent level) and between Groups *C* and *B* (5 per cent level), but not between Groups *C* and *E*.

In summary, hyperventilation, as a means of inducing physiological stress, brings out two clear differences among the "clinical" groups: (*a*) incidence of hypersynchrony or slow waves, and (*b*) degree of heart rate increase. In both instances a tendency toward less adaptive responses is seen in the experimental as compared with the control group.

3. The Cold Pressor Test

The effects of cold pressor stimulation are shown in Table 3 as mean differences between a 60-second period preceding immersion of the hand and

TABLE 3
MEAN RESPONSES TO THE COLD PRESSOR TEST: ACTIVITY DURING IMMERSION
MINUS BASAL ACTIVITY

Measure	Group			
	C	S	B	E
Heart Rate, beats per min.	3.7	-2.1	-2.8	3.4
GSR Index	5.7	0.2	0.1	14.4
Respiration, per min.	1.5	0.8	0.5	1.0
Muscle Potential (μV)	15.9	15.3	5.0	26.0
N	9-10	11	9	4-5

60 seconds during which the hand was in ice water. Since *EEG* difference scores for all measures were negligible they have been omitted from this table; they did not differ significantly between groups of subjects. In fact none of the other mean differences shown in the table differentiates significantly between the groups according to statistical criteria. However an examination of the data for individual subjects indicates that the means for each group tend to mask patterns of response which are shown by certain individuals.

a. Individual patterns of response. With respect to heart rate 24

subjects showed either no change or a slight increase during cold stimulation. Eight subjects showed distinct increases in heart rate and six showed distinct decreases. None of the normal control subjects showed heart rate decreases; 40 per cent showed distinct increases and 60 per cent showed only slight increases or no change. In contrast, of 13 behavior problem children (includes one from Group *S* and two from Group *E* who were problem children) none showed increasing heart rate; 9, or 69 per cent, showed no change and 4, or 31 per cent, showed an atypical or decreasing heart rate. Group *E* subjects tended to be more like the normal control group and Group *S* subjects more like the behavior problem group. The results on the cold pressor heart rate changes for the behavior problem children agree fairly well with those of Roberts (12) who found that none of her normal controls showed decreasing heart rate changes whereas 30 per cent of a group of behavior problem children showed a decreasing rate. Thus it appears that a decreasing heart rate response to cold pressor stimulation is an atypical or abnormal response since it occurs only in some of the behavior problem children and occurs in none of the normal control group.

Descriptively, the usual pattern of heart rate for normal subjects is a slight decrease in heart rate followed by an increase above resting level which persists until the end of stimulation. Immediately after withdrawal of the hand from ice water there is a sharp drop, which appears to be compensatory, and finally gradual recovery to initial resting level, usually within two minutes. This general pattern occurs in almost all subjects who show any increase at all above their basal levels. In cases showing decreasing heart rate there is an immediate and uncomplicated drop upon immersion of the hand with slow recovery and frequently up-and-down fluctuations during the recovery period.

In most subjects no marked change in respiratory pattern occurs during the cold pressor test, although there is a tendency to a slight increase in rate and decrease in depth of breathing. Muscle tension, as shown by the electromyogram, increases in most subjects during the cold pressor test; however, there were no significant differences between groups. The effect was most pronounced in those showing an increasing heart rate. There was no correspondence between muscle tension changes during hyperventilation and those during the cold pressor test.

There were significant positive correlations (rank-difference) among heart rate, galvanic skin response, and muscle tension changes during the cold pressor test. These were as follows: *HR* and *GSR* .59 (*SE* .11); *HR* and *EMG* .46 (*SE* .14); *GSR* Index and *EMG* .60 (*SE* .11).

Figures 1 and 2 show sample records typical for individuals with increasing and decreasing heart rate, respectively, in response to the cold pressor test.

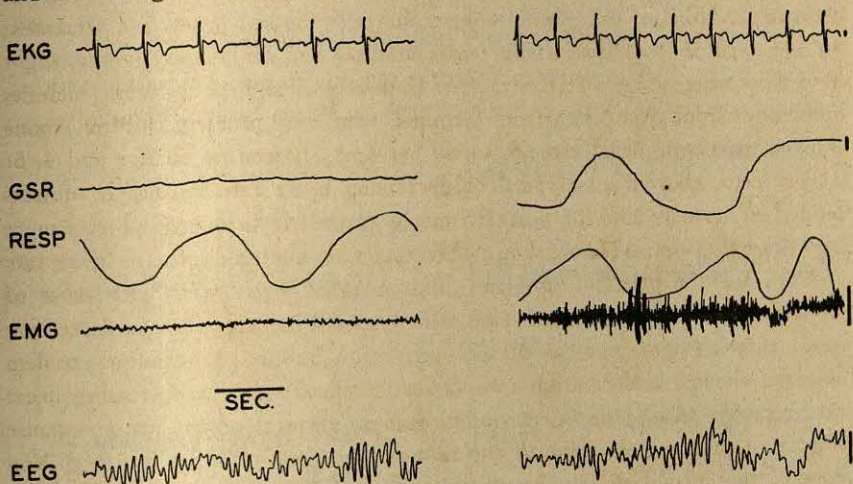


FIGURE 1

Samples of record showing response to the cold pressor test for one control subject. Left: 10 seconds before immersion of the hand in ice water. Right: 10 seconds before removal of the hand from the water. Calibration lines at far right indicate 100 μ V in this and the following figures.

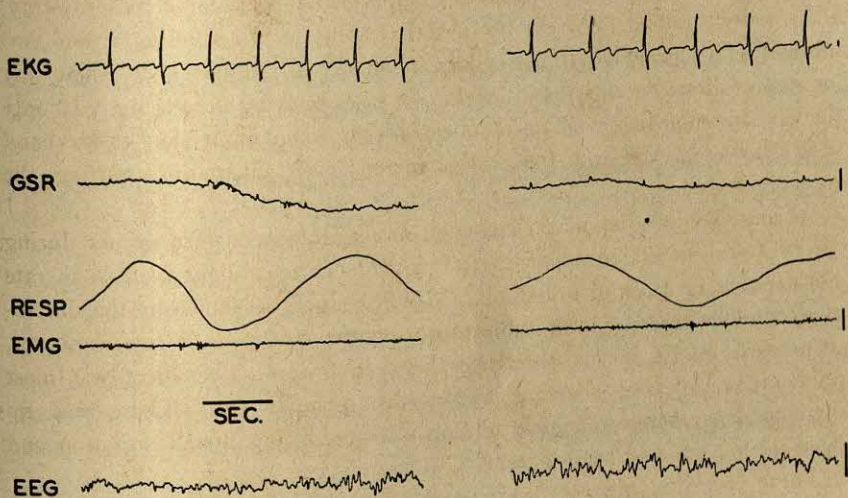


FIGURE 2

Samples of record showing response to the cold pressor test for one behavior problem subject. Left: 10 seconds before immersion of hand. Right: 10 seconds before removal of hand from water.

The former seems to be more characteristic for the normal subjects and the latter for behavior problem children.

4. Response Tendencies of Siblings

Seven pairs of siblings were among the subjects studied. Separate comparisons and correlations between their records have been made. Figure 3 shows

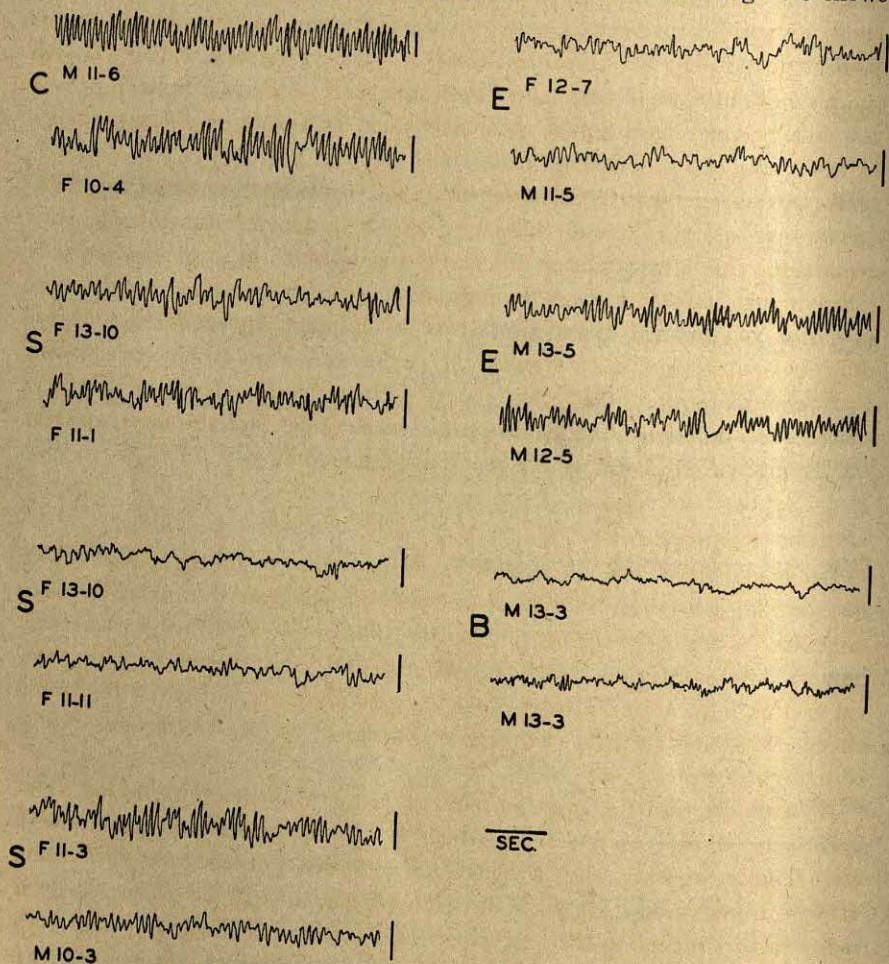


FIGURE 3
EEG tracings from the records of the seven sibling pairs in this study. The similarity of the EEG patterns is immediately apparent. The letters at the left of each pair of tracings indicate the group from which the pair was taken. The figures under each tracing indicate the sex and age (in years and months) of the subject.

the resemblances of the *EEG* patterns for sibling pairs, and quantitative analyses of the records bear this out. For example, rank-difference correlations between the younger and older siblings of each pair and the statistical significance were as follows: percent-time alpha activity, .98 (1 per cent level); percent-time 3 to 8 per second activity, .91 (5 per cent level); delta index, .75 (slightly above 5 per cent level); alpha voltage, .64 (slightly above 5 per cent level). During hyperventilation the correlations were: percent-time alpha, .88 (1 percent level); 3 to 8 per second waves, .70 (not significant); increase in *GSR*, .86 (1 per cent level). During the cold pressor test only two correlations were notable, but not statistically significant: change in heart rate, .68; change in *GSR*, .54.

Two tendencies are to be noted, first, the distinct similarity in the *EEG* patterns of siblings; second, similarity also in autonomic functioning. It is remarkable that any significant correlations were found in so small a group of sibling pairs. However, these findings are in accord with the results of a more extensive investigation of genetic factors in autonomic function (9), and it is difficult to believe that environmental factors can account for such results. Furthermore Kennard (10) has reported close correspondence in the *EEG* patterns of children with behavior disorders and other members of their families; Knott *et al.* (11) have reported similar findings.

E. DISCUSSION

Psychopathologists have become increasingly aware of the difficulties in distinguishing between functional and organic behavior disorders. Probably none are strictly psychogenic, physiogenic, or pathogenic, although there may be very different degrees to which environmental, constitutional, or pathological factors influence each case. Strictly speaking there can be no functional change without structural change, even though it be physico-chemical change of molecular order.

At one extreme lie cases of clearcut structural defect, resulting from genetic or disease processes, in which it may be possible to predict with considerable accuracy that a behavioral disorder or deficiency will ensue. However, the form that such disordered behavior will take depends upon many factors including the past experiences of the individual; but even with a knowledge of these it is not possible to foresee the exact outcome in terms of subsequent symptoms.

At the other extreme, as exemplified in some of the psychoses, there may be no known pathological or physiological alterations concomitant with, or

antecedent to, the behavior disorder. In some of these cases more thorough search has revealed physiological and psychological factors which seem to have importance.

It is with this latter type of problem that this study has been concerned. Behavior problem children, with no known physiological or neurological abnormalities, have been compared, in their physiological reactions, with normal children of the same ages. They have also been compared with two groups of behaviorally normal children, but who might be predisposed to behavior disorders, either by heredity or through early contact with abnormal parents. Physiological reactions during rest and during mild stress conditions were studied.

Several of the behavior problem children differed from the "normals" in a number of physiological reactions and their mode of response appears to be maladaptive and inappropriate to the situation eliciting it. Likewise, behaviorally normal children of schizophrenic parents, as a group, resembled the behavior problem children in all respects in which the latter showed maladaptive responses and differed from the controls. On the other hand behaviorally normal children of epileptics gave responses similar to the normal group, except for hypersynchrony in the *EEG* during hyperventilation.

There is as yet no evidence that deviant physiological responses under stress conditions are related to behavioral aberrations, or whether they may be related as cause or consequence. The most that can be said at present is that the incidence of what appear to be maladaptive physiological responses to stress conditions seems to be greater among behavior children than among normal children. The fact that currently behaviorally normal children of schizophrenic parents tend to be more like behavior problem children with respect to their physiological reactions to stress, and that behaviorally normal children of epileptic parents tend to be more like the normal group raises questions which need further study. Foremost among these are: Do atypical physiological reactions to stress represent a general instability of autonomic and central nervous systems which has predictive value with regard to subsequent behavioral developments and adjustments? Is such an instability hereditary and does it serve as a predisposition to the development of behavior disorders or is it a consequence of psychological maladjustment already begun or experienced, which expresses itself through physiological channels? These and many other questions need intensive and extensive investigation under a variety of physiological and psychological stress conditions, and with adequate follow-up studies of behavioral adjustment over a prolonged period of time.

F. SUMMARY

Four groups of children composed of normals, behavior problems, children with one schizophrenic parent, and children with one epileptic parent have been studied with respect to their physiological reactions to stress conditions induced by hyperventilation and the cold pressor test. Physiological measures consisted of electroencephalogram, electromyogram, galvanic skin response, respiration, and heart rate.

Resting or "basal" levels of activity did not differentiate the groups, but both stress procedures resulted in patterns of somatic and autonomic responses which were differential for normal and behavior problem children. These have been interpreted as instabilities of physiological adjustment, possibly predisposing to behavioral maladjustment. Children of schizophrenics tended to react more like the behavior problem children, whereas the children of epileptics were more like the normal group.

REFERENCES

1. BRILL, N. Q., & SEIDEMANN, H. The electroencephalogram of normal children: Effect of hyperventilation. *Amer. J. Psychiat.*, 1941, **98**, 250-256.
2. DARROW, C. W., & SOLOMON, A. P. Galvanic skin reflex and blood pressure reactions in psychotic states. *Arch. Neurol. & Psychiat., Chicago*, 1934, **32**, 273-299.
3. GOTTLIEB, J. S., KNOTT, J. R., & ASHBY, M. C. Electroencephalographic evaluation of primary behavior disorders in children. *Arch. Neurol. & Psychiat., Chicago*, 1946, **53**, 138-143.
4. HINES, E. A., & BROWN, G. E. A standard test for measuring the variability of blood pressure; its significance as an index of the prehypertensive state. *Ann. Int. Med.*, 1933, **7**, 209-217.
5. HOAGLAND, H., CAMERON, D. E., & RUBIN, M. A. Emotion in man as tested by the delta index of the electroencephalogram. I. *J. Gen. Psychol.*, 1938, **19**, 227-245.
6. HOSKINS, R. G. *The Biology of Schizophrenia*. New York: Norton, 1946.
7. JASPER, H. H., SOLOMON, P., & BRADLEY, C. Electroencephalographic analyses of behavior problem children. *Amer. J. Psychiat.*, 1938, **95**, 641-658.
8. JEFFRESS, L. A. Galvanic phenomena of the skin. *J. Exp. Psychol.*, 1928, **11**, 130-144.
9. JOST, H., & SONTAG, L. W. The genetic factor in autonomic nervous system function. *Psychosom. Med.*, 1944, **6**, 308-310.
10. KENNARD, M. A. Inheritance of electroencephalogram patterns in children with behavior disorders. *Psychosom. Med.*, 1949, **11**, 151-157.
11. KNOTT, J. R., GOTTLIEB, J. S., ASHBY, M. C., & PLATT, E. B. A familial evaluation of the electroencephalograms of patients with primary behavior disorder and psychopathic personality. *EEG & Clin. Neurophysiol.*, 1949, **1**, 249.
12. LINDSLEY, D. B., & CUTTS, K. K. Electroencephalograms of "constitutionally inferior" and behavior problem children. *Arch. Neurol. & Psychiat., Chicago*, 1940, **44**, 1199-1212.

13. MALMO, R. B. Experimental studies of mental patients under stress. In *Feelings and Emotions* (M. L. Reymert, Ed.). New York: McGraw-Hill, 1950. (Pp. 169-180).
14. ROBERTS, D. H. Autonomic activity in normal and behavior problem children. A study of heart rate during physiological and psychological stress. Unpublished Master's Thesis, Brown Univ., 1946.
15. WHITE, B. V., & GILDEA, E. F. "Cold pressor test" in tension and anxiety. *Arch. Neurol. & Psychiat., Chicago*, 1937, **38**, 964-984.

Nebraska Psychiatric Institute
College of Medicine
University of Nebraska
Omaha, Nebraska

THE "NEUROTIC" RORSCHACHS OF NORMAL ADOLESCENTS*¹

Institute of Child Welfare, University of California

VIRGINIA IVES, MARGUERITE Q. GRANT, AND JANE H. RANZONI²

A. INTRODUCTION

In Bell's recent survey of projective techniques (3), the bibliography contains 798 items on the Rorschach, more than eight times as many as are devoted to any other test. The present authors hesitate to add one more item to this list, but hope that in so doing they can hasten the abandonment of a specific approach to the Rorschach which has proved unproductive at least in our research on a cross-section sample.

In this study we have attempted to evaluate the productivity of one widely accepted method of Rorschach analysis, namely, that of "sign patterns." It is concerned with the characteristics of Rorschach protocols commonly accepted by clinicians as "neurotic signs" and "signs of adjustment." This article attempts, via a normal adolescent sample, to throw light on four questions: (a) Do the neurotic signs occur so infrequently or the adjustment ones so frequently in a normal adolescent population that present diagnostic uses are warranted? (b) Does this incidence vary systematically enough over the age span of 11 to 18 to indicate that these Rorschach indices are significantly related to maturity? (c) Can the so-called neurotic and adjustment signs be utilized irrespective of sex? (d) Does the relationship between the number of "adjustment signs" in the Rorschach and ratings of adjustment based upon case histories sustain the claims which have been made for validity?

The Rorschach had its inception in, and its development has been nourished by, clinical psychology and psychiatry. It is not surprising, therefore, that it has been moulded to serve the differential diagnostic needs of the psychiatric clinic, rather than to sharpen projective theory.

*Accepted for publication by Harold E. Jones of the Editorial Board, and received in the Editorial Office on June 11, 1951.

¹Clerical work and statistical analyses were made available by a grant-in-aid from the U. S. Public Health Service. The study proper was made possible by funds from the Laura Spelman Fund, General Education Board, and the Rockefeller Foundation.

²The authors wish to express their appreciation to Dr. Jean W. Macfarlane, Director of the Guidance Study, for her helpful suggestions; to Lucille Allen and Margaret Snyder for administering the tests; and to Rosalie Garber, Ellen Philipsborn, and Karyl Miner for help in the statistical analysis.

Psychological literature reflects this diagnostic concern, in part, in the large number of studies devoted to "sign patterns" on the Rorschach. Such able clinicians as Beck (2), Piotrowski (18), Guirdham (8), Harrower-Erickson (9), and Miale and Harrower-Erickson (16), to name only a few, have perused the Rorschachs of psychiatrically discriminable groups for Rorschach "signs" of so-called disease syndromes.

It is unquestionably important that efforts are bent toward serving the interests of the patient in the psychiatric clinic. However, the value of this diagnostic sign approach to the Rorschach can be questioned in view of the recognition by psychiatry that a diagnostic label is a variable instrument among psychiatrists. The validity, reliability, and usefulness of existing psychiatric diagnostic categories are increasingly questioned (1). Concomitantly, a maturing clinical psychology is now aware that its unique contribution to the understanding of personality does not lie in a slavish devotion to immediate clinical demands and that diagnosis does not guarantee understanding. We should require that our tests contribute to the clarification and development of personality theory as well as to clinical practice.

This paper deals with only two of the many methods of Rorschach analysis, the "neurotic signs" proposed by Miale and Harrower-Erickson (16) and the Davidson signs of adjustment (5). These particular sign lists have not been selected for study because of any characteristics unique to them, but rather as examples of the Rorschach sign approach, to be evaluated through our data on a normal sample. The results of such an evaluation may help to eliminate the misuse of sign patterns which were intended by their authors as signposts rather than goals in themselves.

B. SAMPLE AND PROCEDURES

1. *The Sample*

Our total Rorschach sample consists of 145 adolescents, a sub-sample of the population of the Guidance Study of the Institute of Child Welfare, University of California.³ Comparison with other data indicates that this subsample does not differ significantly from the total sample in any measured way. They cannot be considered representative of the adolescent population as a whole, however, since the community from which the total sample was drawn differs from the average in some socio-economic characteristics. For

³For a complete description of the Guidance Study sample and methods of data collection, see Macfarlane (13). For a more complete description of the Rorschach sample, see McFate and Orr (15).

example, their parents are above the national mean in both education and occupational status.

The sample was divided into two groups for testing purposes. Two-thirds of the group (Series *A*) were first given the Rorschach at age 11, and thereafter at ages 13 and 15. The other one-third of the total sample (Series *B*) were first tested at 12, and thereafter at 14 and 16. The two groups were combined, and all were tested at age 18. (The exact *N*'s for each group and at each age are given in Table 2.) The division was made solely on the basis of birthdate, with the smaller group representing the earliest birthdates. The purpose of splitting the sample was to provide a small separate group for the confirmation of tendencies appearing in the larger group. The small *N* of Series *B* has apparently resulted in greater variation on all measurements at a given age, and less stability of measurements over time. For this reason, we are presenting data from the two series separately, emphasizing the results from the larger *N*.⁴

2. *The Rorschach Procedures*

The sign approach can be and has been called objective, in contrast to the more subjective, Gestalting process of examining the total protocol with its complicated interrelationships in order to make a diagnosis. However, a sign method based on Rorschach scoring can be no more objective than the scoring procedure itself. The methods of administration and scoring employed obviously influenced our results on these signs.⁵ The earliest testing was done in 1939 according to the Klopfer (12) method of administration and scoring. Although there have been changes and revisions of the Klopfer technique since that time, an attempt was made to adhere to the earlier rules throughout the study in order to obtain comparable results at all ages. Wherever known scoring procedures peculiar to the study seem to have influenced the results, this will be pointed out. Over the 10-year period of the Rorschach study, changes in staff necessarily brought new workers whose techniques may have differed slightly. However, all records were conference scored by two or three Rorschach workers, and a test of scoring agreement between the earlier and later scorers gave per cent agreements of .82 to .87 over the age range 11 to 18.

The neurotic sign list to be used in this study was published by Miale and Harrower-Erickson in 1940. (16). This is a list of Rorschach variables

⁴Tables 13 and 14 present the data for Series *B* for the reader's inspection.

⁵For a more complete description of administration and scoring procedures, see McFate and Orr (14).

quantified to test for the presence of neurotic personality structure, and their results and those of a later study by Harrower-Erickson (9) indicate that "scores" on this list, in terms of number of signs, will distinguish between neurotic and non-neurotic groups. Because of the results of her later study, Harrower-Erickson suggested modification of four of the original nine signs and the addition of one new one. In this paper we present data on all 14 of these signs.

The list of signs of adjustment which we have applied to our Rorschach sample was contributed to the literature by Davidson (4). These signs were utilized as measures of adjustment in her study of personality and economic background in intelligent children.

Table 1 lists these neurotic and adjustment signs with definitions of the scoring symbols involved. In order to conserve space, the signs will be referred to hereafter solely in terms of the symbols.

Of the 17 adjustment signs, eight are the converse of items in the "neurotic" sign list. In the signs concerning number of responses, the overlap is such that a subject giving between 21 and 24 responses is scored for *both* an adjustment and a neurotic sign!

Of the signs, color shock and shading shock, as neurotic signs, and no color shock and no shading shock, as adjustment signs, are the most difficult to score objectively. In attempting to assign the presence or absence of these signs to a record, our workers used the 10 criteria for color shock suggested by Brosin and Fromm (4), applying them also to the non-colored cards to test for shading shock as recommended by Klopfer and Kelley (12). As Brosin and Fromm originally presented them, the signs leave operational definition of the criteria to the person who uses them. This flexibility may be desirable for interpreting the color and shading relationships in the individual record, but it is not conducive to reliable work with groups of records. In order to obtain a reasonable level of scoring agreement, it was necessary for the workers to establish rigid and arbitrary limits for the scoring of each of the 10 shock signs. Even this could not be done successfully in all cases.⁶

3. *The External Criteria of Adjustment*

Of the total Guidance Study sample of 252 children, approximately one-half (the *G* group) were seen at yearly intervals from age 6 to age 16 for

⁶The difficulties encountered in an attempt to define these signs in sufficiently objective terms so that they could be reliably scored by more than one worker were so numerous that they, together with a definition of the limits established, will be discussed in a separate paper.

TABLE 1
SYMBOLS AND DEFINITIONS OF NEUROTIC AND ADJUSTMENT SIGNS*

Definition of scoring symbols		Neurotic Sign	Adjustment Sign
R	= number of responses	$R < 25$	$R > 20$
M	= human figures seen in movement	$M < 2$	$M > 2$
FM	= animal figures seen in movement	$FM > M$	$M \leq FM$
Refusals	= refusal to give a response to one or more cards	Refusals	No refusals
F%	= per cent of responses determined by form alone	$F\% > 50$	$F\% \leq 50$
A%	= per cent of animal responses	$A\% > 50$	$A\% \leq 50$
FC	= responses determined by definite form and bright color	$FC < 2$	$FC > 1$
Color shock	= evidence of disturbance on colored cards	CS	No CS
Shading shock	= evidence of disturbance on shaded cards	SS	No SS
FK + Fc	= responses determined by differentiated shading		$FK + Fc > 1$
CF	= responses determined by bright color and indefinite form		$FC \leq CF$
C	= color responses entirely without form		No C
Sum C	= weighted total color responses		$\Sigma C > Fc + c + C'$
$Fc + c + C'$	= shading plus achromatic color responses		$P > 3$ ($< 30\% R$)
P	= popular responses		$\% VIII - X = 40-60$
$\% VIII - X$	= per cent of total responses given to last three cards		$W:M = 2:1$
W : M	= ratio of whole responses to human movement responses		$Dd + S\% \leq 10$
Dd + S%	= per cent of responses to small blot details and/or white space	No FC	
		$R < 12$	
		$FM \leq 2M$	
		No FM	
		$A + At\% > 65$	
A + At%	= per cent of animal plus anatomy responses		

*The first nine neurotic signs are from Miale and Harrower-Erickson (16), and the last five modifications suggested by Harrower-Erickson (9). The adjustment signs are from Davidson (5).

intensive interviews. These interviews, although allowing flexibility, were standard in the sense that the case workers were instructed to cover specific topical areas with each subject in order to provide a basis for quantified ratings of each case. Consequently, the resulting ratings are based on the interviewers' impressions regarding feelings, attitudes, and conflicts as well as on the subjects' verbal responses and those of parents, peers, and teachers.

Included in the ratings based on the interviews are case workers' ratings of the self and social adjustment of the child. The ratings made when the subjects were 11, 13, 15, and 16 years old (the first three corresponding to the ages of Rorschach administrations, and the last, the closest to the 18-year records) were selected for comparisons with the subjects' scores on the Davidson Adjustment Signs.

C. RESULTS

While each of the individual signs has interesting age, sex, and total group characteristics which are available to the reader in the tabled data, the questions raised in the introduction will be given precedence in the discussion of the results. In general, we would like to know: (a) What is the incidence of these signs in a normal adolescent population? (b) Do these Rorschach characteristics show developmental changes? (c) Are there significant sex differences? (d) Do the incidence of the neurotic signs and the correlations of the adjustment signs with interview ratings support the respective "sign" labels? The neurotic signs and adjustment signs will be discussed separately with regard to each of the above questions, and where the adjustment signs results are merely the converse of the findings on comparable neurotic signs, repetition of discussion will be avoided.

1. Incidence

a. *Neurotic signs.* Figure 1 is a graphic presentation of the per cent of our total Rorschach group showing each neurotic sign at each age level, with the sexes graphed separately. Table 2 presents this same material with the addition of data from a "core" group of 37 boys and 34 girls who were given the Rorschach at each of the four age levels.⁷

Even at the older age levels, our group shows a generally higher incidence

⁷The total group consists of all children who were tested at each given age level. The N's for the total group are: 111 at age 11, 122 at age 13, 106 at age 15, and 145 at age 18. However, it was not possible to test every child at every age since the mishaps inherent in a longitudinal study such as illness, change of residence, or resistance to testing sometimes prevented the scheduling of appointments. Hence, the smaller N's for the core group.

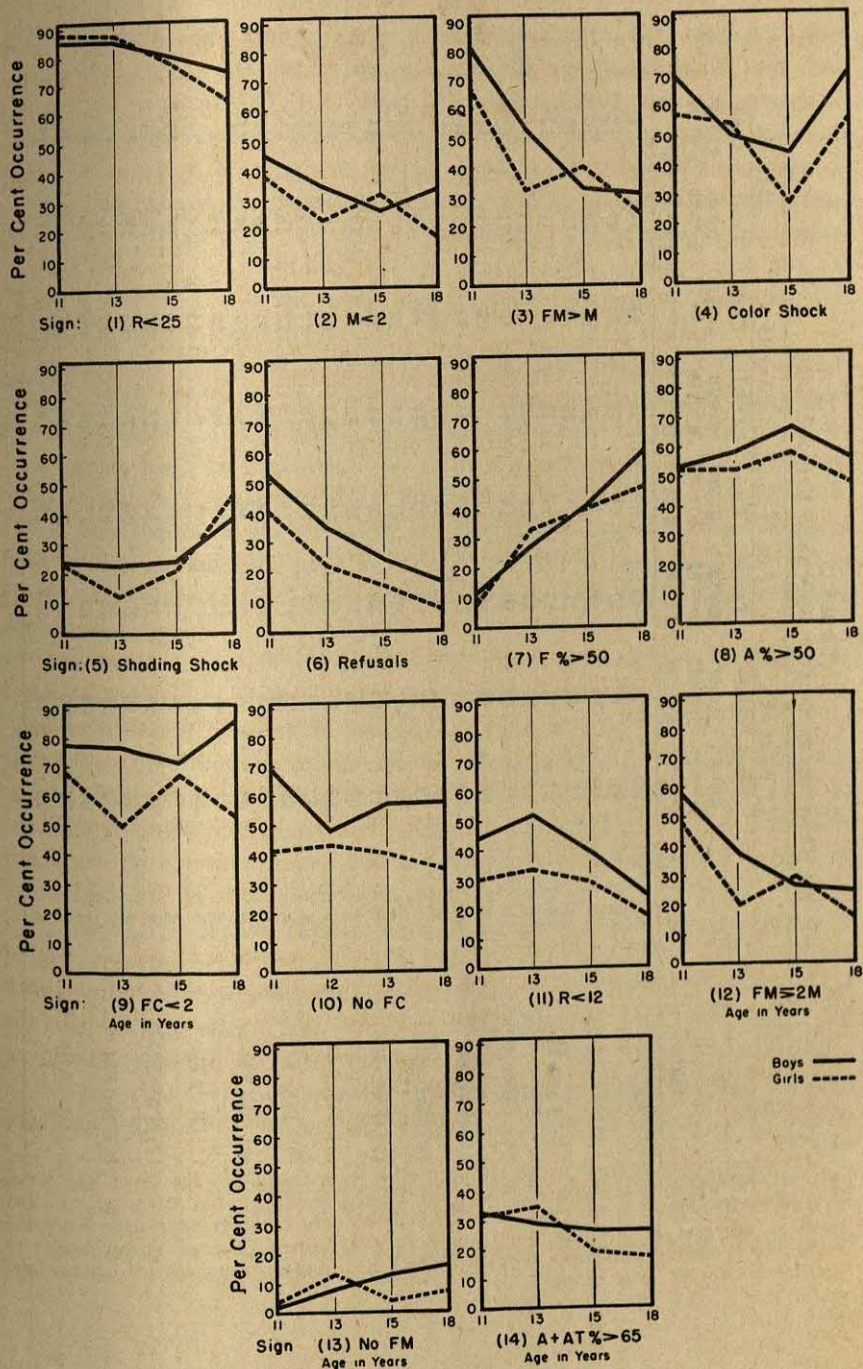


FIGURE 1
PER CENT OCCURRENCE OF RORSCHACH NEUROTIC SIGNS

TABLE 2
PER CENT OCCURRENCE OF RORSCHACH NEUROTIC SIGNS

Sign	Age 11			Age 13			Age 15			Age 18		
	Total group b = 54 g = 53	Core group b = 37 g = 34		Total group b = 63 g = 59	Core group b = 37 g = 34		Total group b = 54 g = 52	Core group b = 37 g = 34		Total group b = 70 g = 76	Core group b = 37 g = 34	
(1) R < 25	85	78	b	85	89		80	84		74	78	
(2) M < 2	88	85	g	87	85		77	76		64	62	
(3) FM > M	45	43	b	35	30		26	27		33	30	
	38	44	g	23	24		31	29		17	18	
(4) Color shock	82	81	b	53	51		33	38		31	32	
	68	76	g	32	29		40	38		24	15	
(5) Shading shock	71	65	b	50	54		44	49		71	65	
	57	59	g	53	59		27	21		55	59	
(6) Refusals	24	24	b	23	24		24	19		39	40	
	23	26	g	12	06		21	24		47	50	
(7) F% > 50	53	57	b	35	32		24	30		16	08	
	41	47	g	22	21		15	18		07	03	
(8) A% > 50	11	14	b	27	27		41	38		59	54	
	07	06	g	33	29		40	38		47	44	
(9) FC < 2	53	54	b	58	46		67	65		56	60	
	52	53	g	52	53		58	59		48	41	
(10) No FC	78	76	b	77	84		72	73		86	84	
	68	68	g	50	68		67	71		53	50	
(11) R < 12	69	68	b	48	54		57	57		57	51	
	41	41	g	43	47		40	38		35	29	
(12) FM \geq 2M	44	43	b	52	57		39	43		24	19	
	30	35	g	33	32		29	24		17	12	
(13) No FM	58	54	b	37	38		26	27		24	24	
	48	56	g	20	24		29	26		15	09	
(14) A + At% > 65	02	03	b	08	08		13	11		16	14	
	04	00	g	13	15		04	06		07	06	
	33	30	b	29	24		26	27		26	24	
	32	38	g	35	15		19	15		17	15	

of the neurotic signs than might be expected in a normal population if these signs were truly indicative of neurosis. In the interpretation of such data, there are two possible points of view. We might proceed upon the assumption that these signs are valid indicators of neurosis and so conclude that our adolescents are, as a group, not well adjusted to themselves or their environments. Or we might assume that a large random sample of the adolescent population will represent an adequate sampling of "adjustment" in a psychologically meaningful sense and that the most fruitful approach is to test the signs against the population rather than the population against the signs. The authors prefer the latter approach.

Particularly interesting in terms of incidence are the signs $R < 25$, $FC < 2$, and No FC. Although there are no exactly comparable groups with regard to both sample and Rorschach procedures, the mean number of responses for our sample has been consistently lower than most reported in the literature. Our R may be at least partially a function of the method of administration, since our Rorschach workers made no attempt to prod for more responses than were spontaneously given by the subject. When the sign is modified to $R < 12$, the incidence (24%) approaches more nearly the level to be anticipated in a normal group.

The neurotic signs $FC < 2$ and No FC, which are related to lack of color responsiveness, also show a surprisingly high incidence in our population. Approximately 40 per cent of the girls and more than 50 per cent of the boys at each age level have no color responses. Regardless of the interpretation one might make of this lack of color responsiveness in our population, it is one of our most significant findings and is consistent through all of our Rorschach analyses. At the same time, it is not consistent with some other reports of Rorschach studies of adolescence in the literature.⁸

Table 3 gives the comparative data on incidence of the individual signs in groups reported in the literature (10 and 12) and comparable age levels from our groups. We are using the data from Harrower-Erickson's second report on the neurotic signs (10), since, as she points out, both the normal and neurotic groups are larger and more heterogeneous than her previous sample.

In terms of incidence, our 18-year-olds seem to stand midway between Harrower-Erickson's normals and neurotics. They are significantly⁹ higher

⁸Hertz (11), for example, reports a much higher mean color response for her adolescents than our mean FC, which ranges from 0.71 to 1.84 for boys and girls over our total age range.

⁹The significance of the differences in percentage incidence was tested by means of the tables and formula of Edgerton and Patterson (7).

TABLE 3
COMPARISON OF GUIDANCE STUDY GROUPS WITH DATA FROM HARROWER-ERICKSON AND HERTZMAN AND MARGULIES*

Sign	Boys			Boys			Boys and Girls			Boys and Girls		
	G.S. 13 N = 63	H-M N = 60	Jr. Hi. N = 60	G.S. 18 N = 70	H-M N = 62	Coll. N = 62	G.S. 18 N = 145	H-E N = 145	Norm. N = 385	G.S. 18 N = 145	H-E N = 145	Neur. N = 74
	% Incidence	% Incidence	% Incidence	% Incidence	% Incidence	% Incidence	% Incidence	% Incidence	% Incidence	% Incidence	% Incidence	% Incidence
(1) R < 25	85	75		74	15		69	62		69	97	
(2) M < 2	35	47		33	18		25	32		25	72	
(3) FM > M	53	60		31	73		28	57		28	86	
(4) Color shock	50	58		71	56		63	6		63	46	
(5) Shading shock	23	43		39	32		43	5		43	59	
(6) Refusals	35	23		16	8		11	12		11	65	
(7) F% > 50	27	53		59	34		52	19		52	57	
(8) A% > 50	58	67		56	27		**					
(9) FC < 2	77	85		86	31		***45	29		45	79	

*Underlined differences are significant at the one per cent level of confidence.

**In the comparisons with the Harrower-Erickson groups, Sign 8 cannot be used since she scored either A% > 50 or A + At% > 50 in her second study.

***These comparisons are based on No FC rather than FC < 2, as Harrower-Erickson used the former in her second study.

than her normals in four of the eight comparable signs, significantly lower on one, and similar on the others. On the other hand, they are significantly lower than Harrower-Erickson's neurotics on six of the eight signs, significantly higher on one, and equal on the sign of $F\% > 50$.

Even after establishing extremely strict and limiting criteria for assigning color shock, our normal group at 18 shows a higher incidence than the Harrower-Erickson neurotics on this sign. It is also worthy of note that, although they are lower than the Harrower-Erickson neurotic group as a whole, five of the eight signs showed a frequency of 40-70 per cent in our 18-year-old group.

In the study by Hertzman and Margulies (12), the neurotic signs were tested for age changes from a group of junior high school boys to a matched group of male college students. (The groups were matched for intelligence, academic status, and socio-economic status.) We have also compared our 13-year-old boys with Hertzman and Margulies' junior high school students and our 18-year-old boys with their college students. As Table 3 shows, sign incidence in the younger groups is quite similar, with only one significant difference. However, there is a consistent tendency for our 18-year-olds to show a higher incidence than their college students on all but one of the nine signs. This may be related to the highly selected nature of the Hertzman and Margulies' subjects (Jewish New York students of superior intelligence) or a result of the age difference between the two groups. (Hertzman and Margulies' college students had a mean age of 19.4.) Here again we would like to point out, however, that our group has over 40 per cent incidence of five of the nine neurotic signs, while the Hertzman-Margulies group is below 40 per cent on seven of them.

b. *Adjustment signs.*¹⁰ Figure 2 presents graphically the per cent incidence of each of the adjustment signs for both sexes at each of the four age levels in Series A. In Table 4 these incidence data are given for the total group and for the "core" group.

As Figure 2 illustrates, the incidence of the adjustment signs varies considerably from one sign to another. The sign, No C, occurs in 90 per cent or more of the sample at all ages. The signs, $\text{Sum } C > Fc + c + C'$ and $W:M$

¹⁰Since the completion of the present data analysis, another paper by Davidson has appeared in the *Journal of Projective Techniques* (6), in which she summarizes data from published and unpublished studies relative to the adjustment signs. She particularly notes the need for further studies to provide evidence on the reliability and validity of the signs. All of the material on adjustment signs to follow throws some light on the questions raised by Davidson.

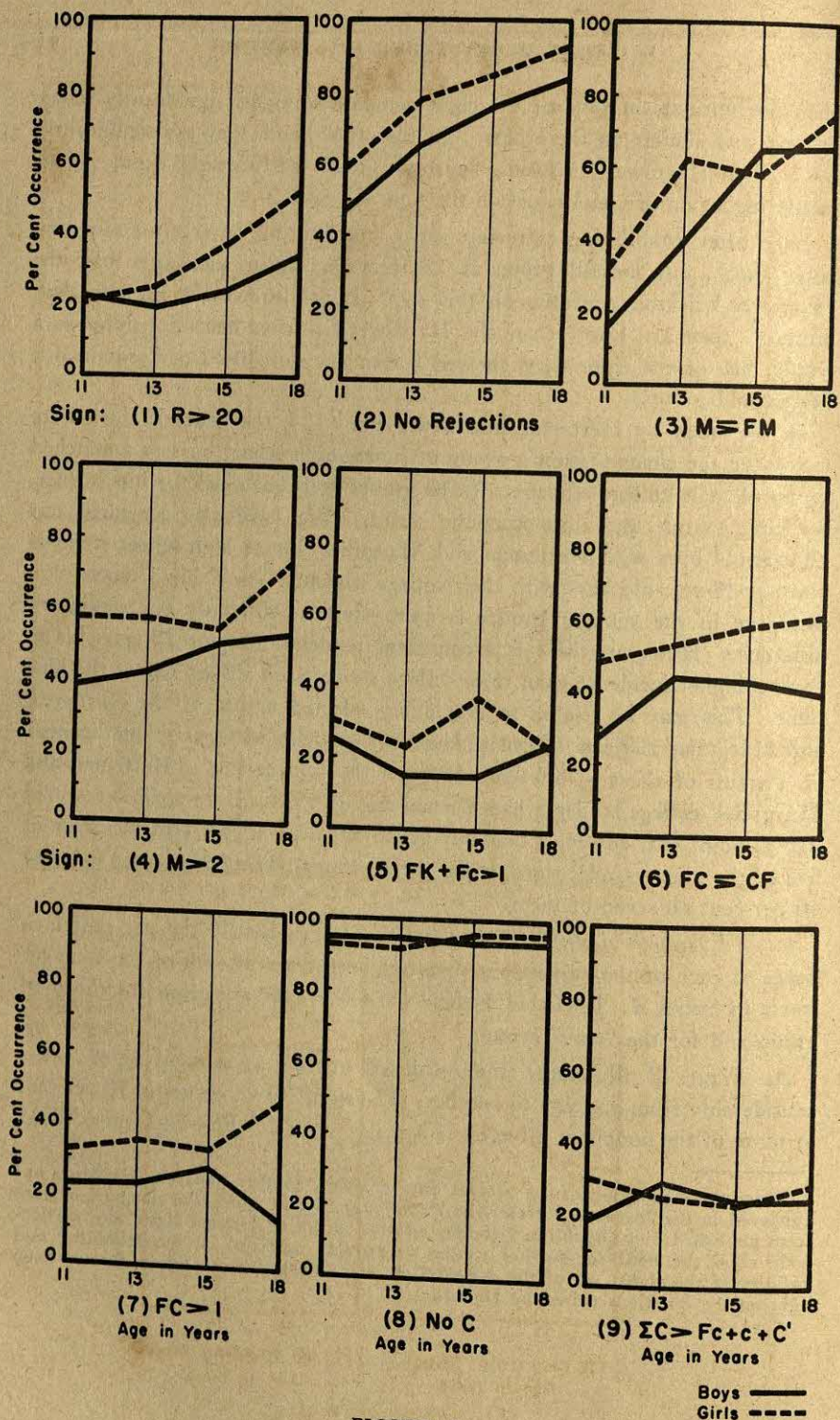


FIGURE 2
PER CENT OCCURRENCE OF RORSCHACH ADJUSTMENT SIGNS

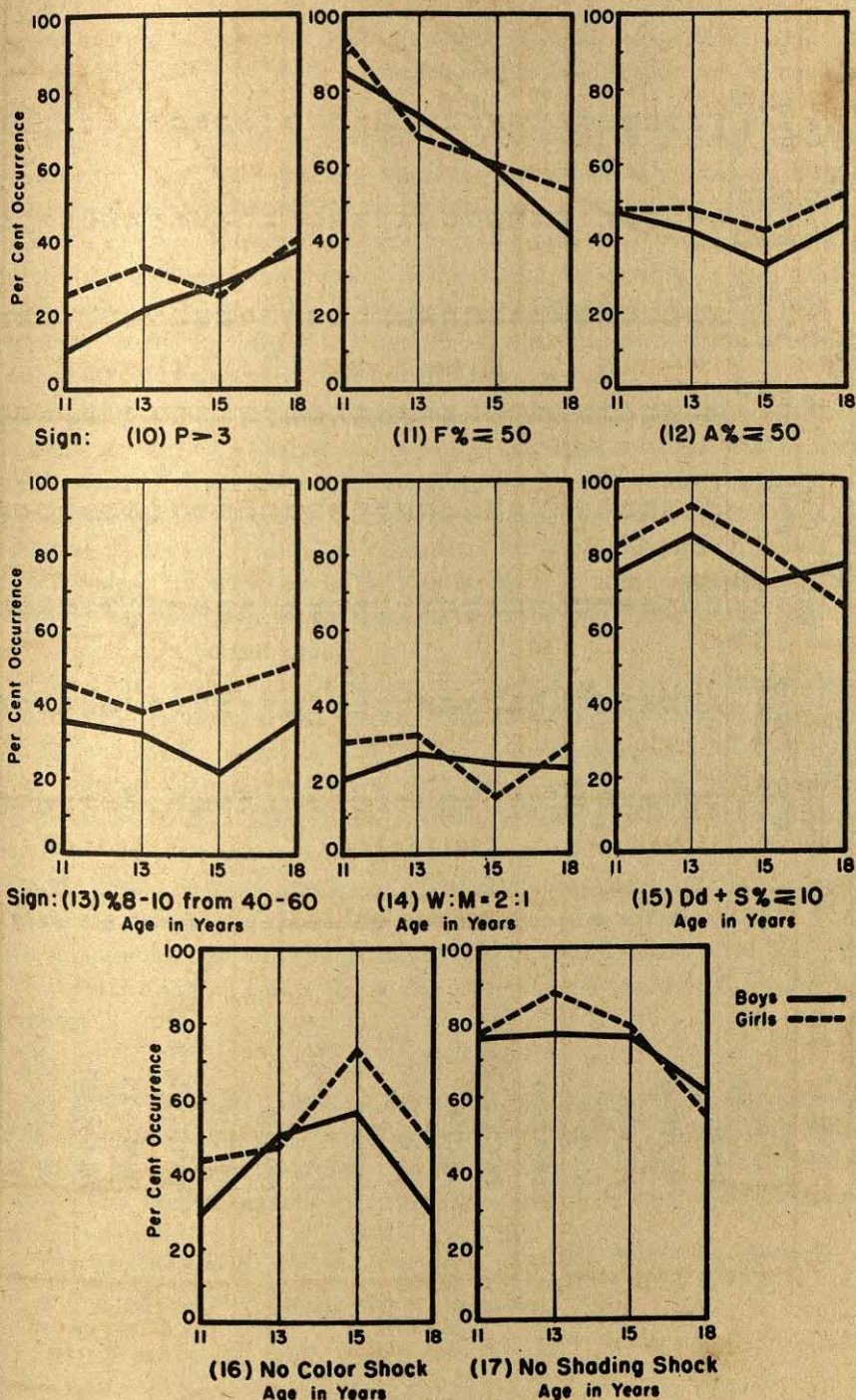


FIGURE 2 (continued)

TABLE 4
PER CENT OCCURRENCE OF RORSCHACH ADJUSTMENT SIGNS

Sign	Age 11			Age 13			Age 15			Age 18		
	Total group b = 54 g = 53	Core group b = 37 g = 34		Total group b = 63 g = 59	Core group b = 37 g = 34		Total group b = 54 g = 52	Core group b = 37 g = 34		Total group b = 70 g = 76	Core group b = 37 g = 34	
R > 20	22	24	b	19	16		24	22		34	30	
No refusals	21	24	g	25	24		37	38		52	56	
	47	43	b	65	68		76	70		84	65	
M \leq FM	59	53	g	78	79		85	82		93	97	
	16	16	b	40	43		65	62		66	68	
M > 2	32	24	g	62	62		58	59		75	82	
	38	38	b	42	46		50	49		53	51	
FK + Fc > 1	57	47	g	57	50		54	50		73	79	
	25	22	b	15	16		15	14		24	19	
FC \leq CF	30	27	g	23	29		37	44		23	24	
	27	32	b	44	38		43	43		23	24	
FC > 1	48	44	g	53	50		58	62		61	71	
	22	24	b	23	16		28	27		14	16	
No C	32	32	g	35	32		33	29		47	50	
	95	97	b	95	92		94	97		94	95	
2C > Fc + c + C'	93	94	g	92	85		96	94		96	100	
	18	14	b	29	32		24	19		24	24	
P > 3 (< 30% R)	30	26	g	25	18		23	18		28	18	
	16	19	b	21	19		28	24		37	32	
F% \leq 50	25	32	g	33	32		25	24		40	41	
	85	86	b	73	73		59	62		41	46	
A% \leq 50	93	94	g	67	71		60	62		53	56	
	47	46	b	42	54		33	35		44	41	
% VIII-X is 40-60	48	47	g	48	47		42	41		52	59	
	35	38	b	32	32		22	16		36	51	
W:M = 2:1	45	53	g	38	44		44	53		51	56	
	20	22	b	27	24		24	24		23	27	
Dd + S% \leq 10	30	24	g	32	24		15	18		29	26	
	75	76	b	85	92		72	76		77	81	
No color shock	82	82	g	93	97		81	85		65	74	
	29	35	b	50	46		56	57		29	32	
No shading shock	43	41	g	47	41		73	79		47	44	
	76	76	b	77	76		76	81		61	59	
	77	74	g	88	94		79	76		55	53	

approximately 2:1, occur in less than one-third of the group at all ages. Of the 17 signs, only nine reach an incidence of at least 60 per cent at any age for either sex.

While it is difficult to analyze the 17 signs as a whole, it seems justifiable to state that they occur with less frequency than would be anticipated if they were validly representative of the normal or "adjusted" Rorschach. The selection of items for the adjustment sign list was based on "the experience of Rorschach investigators, and available norms for healthy, normal adults and children" (5, pg. 88). At the time of the Davidson study, and even at present, "available norms" are meager and unsystematic, often presented informally and based upon biased or incomplete sampling.

In addition, the Davidson signs were validated only by the method of internal consistency. That is, each sign was correlated with the total number of signs and, with one exception, signs correlating $+ .23$ or more with the total were retained. The majority of research workers would now agree that some external criterion is essential to validation. For these reasons and because of other evidence of weakness in the sign methods, it is not surprising to find that the sign list as a whole does not seem to be related to the level of adjustment of a normal group.

The sample with which Davidson utilized her signs of adjustment consisted of 102 children between the ages of 9 and 13. Their *IQ*'s ranged from 125 to 200 plus, with a mean of 142.6. Our Rorschach subjects are 11 to 18 years old, with an *IQ* range of 60 to 180 and a mean *IQ* of 119 (standard deviation = 16.28). It is not our intention to compare our results with those of Davidson since the samples and methods are not comparable. However, the one datum of mean number of adjustment signs may be of some interest. The mean number of signs for the Davidson group was 10.31. Table 5 gives the mean number of adjustment signs for both sexes at each of the four age levels studied in our group. In no case does it reach the level of Davidson's findings.¹¹

2. Age Changes

a. Neurotic signs. In their study of age differences in the neurotic signs in matched groups, Hertzman and Margulies (12) found their older group to have a lower incidence than the younger on all but one of the nine signs. This finding suggests that these signs may be related to chronological

¹¹The correlation of *IQ* and total number of adjustment signs is not significant in our group. The Pearsonian *r*'s are $-.01$ at 11 years, $+.06$ at 13, $+.04$ at 15, and $+.18$ at 18. Davidson (5) reports $+.24$ for her group.

TABLE 5
MEAN NUMBER OF ADJUSTMENT SIGNS

Age	Boys	Girls	Total
11	6.98	8.46	7.73
13	7.79	8.97	8.37
15	7.89	8.98	8.42
18	7.81	9.40	8.63

(and, presumably therefore, emotional) maturity. Such a relationship would be entirely consistent with the concept of "neurotic" signs, since emotional immaturity plays a major role in neurosis. However, our data do not support this hypothesis.

Our data offer an opportunity to test the developmental characteristics of the neurotic signs in the *same* children over the crucial period of adolescence. The sign incidence in the core group (Table 2) was tested for the significance of age changes. Table 6 gives the level of significance of the critical ratios obtained at each age level.¹²

Obviously these data do not show any consistent age trends over all of the 14 signs tabled. A few of the signs show significant decreases with age. Some show no change at all, while others show significant increases with age. There is not enough consistency in the changes to assume that any of the signs individually (with the possible exception of FM > M, Refusals, and FM \leq 2M) nor the sign list as a whole are related to age.

One of the most startling results in this respect for our group is the neurotic sign F% > 50, which increases consistently and significantly with age until, at 18, 50 per cent of the boys and 47 per cent of the girls show this sign. In the absence of confirmation in the literature (Hertzman and Margulies (12) found their older group to have a significantly lower incidence than the younger group), one is tempted to advance some interpretive hypotheses concerning the characteristics of our group. (For example, the permissive atmosphere of Berkeley and the west coast in general may make greater internal structuring necessary for these adolescents. In addition, the decrease in FM in our group along with the increase of F%, both findings the converse of Hertzman and Margulies' results, suggest the dropping out of immature impulse patterns without the development of mature patterns to replace them.) However, at this stage of the analysis of our Rorschach

¹²The critical ratio formula is from McNemar (16), pg. 80, formula 28. It takes account within the method of computation of the correlational factor involved in spaced testing of the same individuals.

TABLE 6
SIGNIFICANT AGE CHANGES IN INCIDENCE OF RORSCHACH NEUROTIC SIGNS

Sign		11-13	11-15	11-18	13-15	13-18	15-18
1. $R < 25$	b	-	-	-	-	5	-
	g	-	-	1	-	1	5
2. $M < 2$	b	-	-	-	-	-	-
	g	5	-	1	-	5	5
3. $FM > M$	b	1	1	1	-	-	-
	g	1	1	1	-	5	5
4. Color shock	b	-	-	-	-	-	(5)
	g	-	1	-	1	-	(1)
5. Shading shock	b	-	-	-	-	-	-
	g	5	-	(5)	(5)	(1)	(5)
6. Refusals	b	5	5	1	-	1	1
	g	1	1	1	-	1	5
7. $F\% > 50$	b	-	(5)	(1)	-	(1)	-
	g	(1)	(1)	(1)	-	-	-
8. $A\% > 50$	b	-	-	-	(5)	-	-
	g	-	-	-	-	-	-
9. $FC < 2$	b	-	-	-	-	-	-
	g	-	-	-	-	-	-
10. No FC	b	-	-	-	-	-	-
	g	-	-	-	-	-	-
11. $R < 12$	b	-	-	5	-	1	1
	g	-	-	5	-	5	-
12. $FM \leq 2M$	b	-	1	1	-	-	-
	g	1	1	1	-	5	5
13. No FM	b	-	-	(5)	-	-	-
	g	(5)	-	-	-	-	-
14. $A + At\% > 65$	b	-	-	-	-	-	-
	g	5	5	5	-	-	-

1 = 1% level of confidence.

5 = 5% level of confidence.

- = Non-significant.

() = Significant increase rather than decrease with age.

data, we wish to avoid interpretation, leaving that to later reports which will integrate our projective findings with other clinical and measurement data.

The only signs which show no change at all with age are the two ($FC < 2$ and No FC) which are related to color responsiveness. Hertzman and Margulies (12) report one of their largest differences on the sign $FC < 2$, with a much smaller percentage of their older group showing the sign.

Of the three signs which show most consistent decreases with age in our group, two ($FM > M$ and $FM \leq 2M$) concern the M to FM ratio, which has always been considered fundamental in the estimation of maturity level from the Rorschach.

Overlapping the two topics of age change and of sex difference (to be presented in the next section) is the very obvious sex difference in age

changes. The girls show more than twice as many significant decreases with age as the boys. The sex differences in maturity levels during adolescence have already been well established in the research literature. However, the consistency of our findings over a wide age range which approaches adulthood suggests that sex differences in these Rorschach signs deserve more attention than they have received. The fact that there are as many significant sex differences at 18 as there are at 11 and 13 (see Table 9) suggests that more than maturity is involved.

b. Adjustment signs. It is not entirely clear from the Davidson monograph whether these signs were chosen as particularly applicable to the pre-adolescent and adolescent group with which she worked. However, all of them are essentially in line with the suggested norms for adult records as presented by Klopfer and Kelley (13). On the other hand, most reliable work with the Rorschachs of children and adolescents takes account of the differences to be expected from adult norms. Since the signs do correspond to published norms for adult records, one might expect an increase in their incidence as the age of the group approaches adulthood.

The per cent incidence of each of the adjustment signs in our sample was tested for age changes over the seven-year range. Table 7 summarizes the significant age changes.

Only two of the signs (No Refusals and $M \geq FM$) show consistent increases of incidence with age for both sexes. For girls, the signs $R > 20$ and $M > 2$ seem to increase significantly with age. The majority of the signs show no consistent developmental trends. The signs $F\% \leq 50$, for both sexes, and No Shading Shock, for girls, show a definite decrease in incidence from age 11 to age 18 (as their opposites in the neurotic signs showed increases).

Here again it is difficult to interpret the findings in terms of the sign list as a whole. Surprisingly few of the signs which are considered normal in adult records seem to increase with age during adolescence. In the 11-18-year comparison, covering the entire age range for our group, the only signs which show evidence of change in either sex are related to the very basic Rorschach variables of number of responses, movement balance, and color balance. The same was true of the neurotic signs.

While Table 7 concerns age changes in incidence of each of the *individual* adjustment signs, Table 8 gives the percentage of children who showed an increase in *total number of adjustment signs* over the two- or three-year intervals between tests, the percentage who showed a decrease, and the

TABLE 7
SIGNIFICANT AGE CHANGES IN INCIDENCE OF RORSCHACH ADJUSTMENT SIGNS

Sign		11-13	11-15	11-18	13-15	13-18	15-18
1. $R > 20$	b	-	-	-	-	5	-
	g	-	-	1	5	1	5
2. No refusals	b	5	1	1	-	1	(1)
	g	1	1	1	-	1	5
3. $M \equiv FM$	b	5	1	1	-	5	-
	g	1	1	1	-	5	5
4. $M > 2$	b	-	-	-	-	-	-
	g	-	-	1	-	1	1
5. $FK + Fc > 1$	b	-	-	-	-	-	-
	g	-	-	-	-	-	-
6. $FC \equiv CF$	b	-	-	-	-	-	-
	g	-	-	5	-	5	-
7. $FC > 1$	b	-	-	-	-	-	-
	g	-	-	-	-	-	-
8. No C	b	-	-	-	-	-	-
	g	-	-	-	-	5	-
9. $\Sigma C > Fc + c + C'$	b	5	-	-	-	-	-
	g	-	-	-	-	-	-
10. $P > 3$ ($< 30\%R$)	b	-	-	-	-	-	-
	g	-	-	-	-	-	5
11. $F\% \equiv 50$	b	-	(5)	(1)	-	(1)	-
	g	(1)	(1)	(1)	-	-	-
12. $A\% \equiv 50$	b	-	-	-	(5)	-	-
	g	-	-	-	-	-	-
13. %VIII-X is 40-60	b	-	-	-	-	-	1
	g	-	-	-	-	-	-
14. $W:M = 2:1$	b	-	-	-	-	-	-
	g	-	-	-	-	-	-
15. $Dd + S\% \equiv 10$	b	1	-	-	(5)	-	-
	g	5	-	-	(5)	(1)	-
16. No color shock	b	-	5	-	-	-	(5)
	g	-	1	-	1	-	(1)
17. No shading shock	b	-	-	-	-	-	-
	g	5	-	(5)	(5)	(1)	(5)

1 = 1% level of confidence.

5 = 5% level of confidence.

- = Non-significant.

() = Significant increase rather than decrease with age.

percentage whose total number of adjustment signs remained the same. These data, too, show an absence of developmental trends. The per cent of subjects showing an increase in total number of signs is only slightly larger in most of the age comparisons than those showing a decrease. In the last two rows of Table 8, which compare the per cent increasing with the per cent not increasing, there is no significant difference except in the 13-15 comparison where the per cent not showing an increase is larger.

TABLE 8
PER CENT OF GROUP SHOWING AN INCREASED, DECREASED, OR EQUAL
NUMBER OF ADJUSTMENT SIGNS FROM AGE TO AGE

	11-13	13-15	15-18	11-18
(A) Per cent showing increase				
Per cent of boys	53	35	40	51
Per cent of girls	56	44	60	66
Per cent of total group	55	39	50	59
(B) Per cent showing decrease				
Per cent of boys	30	44	43	29
Per cent of girls	41	28	31	26
Per cent of total group	35	37	37	27
(C) Per cent showing no change				
Per cent of boys	16	21	17	20
Per cent of girls	3	28	9	9
Per cent of total group	10	24	13	14
Total per cent increasing (A)	55	39	50	59
Total per cent not increasing (B+C)	45	61*	50	41

*Difference between (A) and (B+C) significant at 1 per cent level of confidence.

3. Sex Differences

a. *Neurotic signs.* All of the significant sex differences favor the girls; that is, a smaller percentage of the girls show the neurotic signs. Table 9 summarizes the significant sex differences.¹³

Of these differences, all but one ($R < 12$) are related to the very basic Rorschach variables of movement and color, and emphasize the greater need for the recognition of sex differences in all work with this technique.

Table 9 shows no sex differences at age 15. This is of particular interest since an examination of Figure 1 shows how frequently the curve trend is broken at 15, and Table 6 indicates how few age changes in the signs occur

TABLE 9
SIGNIFICANT SEX DIFFERENCES IN RORSCHACH NEUROTIC SIGNS

Sign	Age	Level of confidence	Smaller % incidence
(2) $M < 2$	18	.05	Girls
(3) $FM > M$	13	.05	Girls
(4) Color shock	18	.05	Girls
(9) $FC < 2$	13 and 18	.01	Girls
(10) No FC	11 and 18	.01	Girls
(11) $R < 12$	13	.05	Girls
(12) $FM \leq 2M$	13	.05	Girls

¹³Sex differences in per cent incidence were computed by means of the tables and formula of Edgerton and Patterson (7).

in the 13-15 comparison. The meaning of these deviant results at age 15 will only become clear when related to data from other sources.

b. Adjustment signs. The sex characteristics of the adjustment signs are similar to those of the neurotic signs. Table 10 summarizes the significant sex differences, all of which favor the girls (a higher percentage of girls show the adjustment signs). In addition to those signs which duplicate the neurotic signs, the signs $FC \leq CF$ and %VIII-X is 40-60, both related to color responsiveness, show a significant sex difference.

TABLE 10
SIGNIFICANT SEX DIFFERENCES IN RORSCHACH ADJUSTMENT SIGNS

Sign	Age	Level of confidence	Sex showing highest incidence
(1) $R > 20$	18	.05	Girls
(3) $M \leq FM$	11 and 13	.05	Girls
(4) $M > 2$	11 and 18	.05	Girls
(6) $FC \leq CF$	11 and 18	.05 and .01	Girls
(7) $FC > 1$	18	.01	Girls
(13) %VIII-X is 40-60	15	.05	Girls
(16) No color shock	18	.05	Girls

The implications for interpretation of these sex differences on the Rorschach cannot be ignored. However, we are aware of one fact which may influence the magnitude and consistency of the sex differences in our group. At the time of the later Rorschach administrations, our boys were faced with the prospect of the draft and military service, which may account in some part for the fact that their Rorschachs seem more disturbed. However, this was not imminent when the boys were 11 and 13, where sex differences also occur.

4. The Lists as "Signs"

a. Neurotic signs. We have raised the question of whether or not our findings will support the label of "neurotic sign list." In general, we must answer in the negative. In summarizing our conclusions, however, we must also attempt to justify the drawing of these conclusions from a sample such as ours.

Perhaps the most vulnerable characteristic of our sample for the purpose of testing the neurotic signs is the fact that it is a population of adolescents rather than adults. A merely casual perusal of these data might lead to the conclusion that one would expect a higher incidence of these signs in adolescents than in adults, greater sex differences, and a change in the

incidence of the signs with age, and that these expectations are borne out. Can anything, then, be said about their applicability as neurotic signs in adult records? We feel that it can. Our 18-year-olds show an incidence of 40-86 per cent on seven of the signs. Assuming that they can be treated as adult records (the use of college freshmen and sophomores for adult norms has a respectable precedent), this incidence is considerably greater than expectation in a normal group. Sex differences are as frequent and as great at 18 as they are at 13. And certainly we do not find enough evidence of consistent decrease of these signs with age to assume that they are related to chronological age, and that, had we obtained a sample of adults, we would find a significantly lower frequency of the neurotic signs in a normal population.

Our sample has a specific advantage for this purpose in that it is a fairly random population sample. It was not selected to test differences already known to exist (i.e., hospitalized vs. non-hospitalized, psychosomatic vs. somatic, etc.). We are not concerned, therefore, with the need to discriminate between groups. Harrower-Erickson (10 and 17) suggests five neurotic signs as a cut-off point for the diagnosis of neurosis. Figure 3 shows the

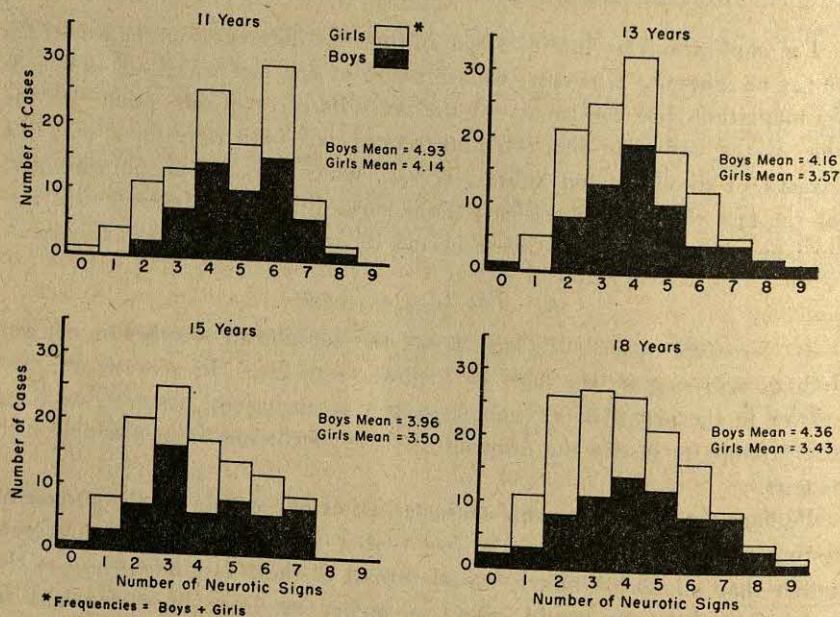


FIGURE 3
DISTRIBUTION OF NUMBER OF NEUROTIC SIGNS OVER AGES, TOTAL GROUP

distribution of number of neurotic signs for our group. The means at all ages are close to this cut-off point. At age 11, 51 per cent of our group have five or more signs, 31 per cent at 13, 32 per cent at 15, and 36 per cent at 18. This last figure is quite at odds with the reported 6 per cent of Harrower-Erickson's 154 male and female college students who had five or more signs (10).

Certainly on the basis of our results, we would advise against the use of these signs for discriminating between individual Rorschach records as neurotic or non-neurotic. This is especially true as we become more aware through experience that distinctions between "normal," "neurotic," and "psychotic" are difficult to make even on the basis of extensive clinical material and contacts. In addition, the magnitude of the critical ratios obtained between groups in Rorschach research has never guaranteed much more than chance ability to distinguish individuals. In view of these facts and on the basis of our data, it would seem that the diagnostic value for individuals of some current approaches to test material should be questioned.

As we have pointed out, more than one-third of our adolescent group would be labeled "neurotic" on the basis of showing the critical number of signs. Had we accepted the single fundamental interpretive variables upon which the list was based, and proceeded from them to analyze our group, we would find that at age 18, 69 per cent¹⁴ of our adolescents are underproductive ($R < 25$), 52 per cent are immature in their thinking ($A\% > 50$), 52 per cent are rigid and constricted ($F\% > 50$), 69 per cent are emotionally unresponsive ($FC < 2$), 46 per cent completely so (No FC), 63 per cent are disturbed by strong emotional stimuli (Color Shock), and 43 per cent by stimulation of their affectional needs (Shading Shock). The life adaptation of these 18-year-olds precludes such interpretation.

This is, of course, an oversimplification of our data. It is used, however, to point up the necessity for a more comprehensive handling of Rorschach material. To some extent these things may be true of adolescents. If they are true, however, the signs have little value in discriminating between normal and neurotic subjects.

b. Adjustment signs. Thus far we have interpreted the results of the sign study solely in terms of incidence in our normal sample. We have questioned the validity of the sign lists as a whole on the basis of these group data. It is possible, however, to relate our findings on the adjustment signs to ratings of adjustment by case workers as an external criterion.

Since the study at the Institute of Child Welfare has been carried on with a

¹⁴All per cents refer to the total group, boys and girls combined.

normal group of children, the interviewing and descriptive measurement has been directed toward concepts of normality rather than pathology. While no evaluation of the "neurotic" status as such is available in our sample, ratings of self and social adjustment by case workers for the individual child have been made and can be compared with the Rorschach analysis.¹⁵

Davidson has divided her 17 signs into six indicating personal or self adjustment (Signs 3, 4, 9, 10, 11, and 15 as we list them) and the other 11 indicating social adjustment. We have correlated "scores" (in terms of number of signs) on these parts of the list with interview ratings of self and social adjustment. In addition, we have correlated total "score" on the 17 signs with the combined self and social ratings from the interview. These correlations are presented in Table 11.

TABLE 11
CORRELATION OF NUMBER OF DAVIDSON ADJUSTMENT SIGNS WITH RATINGS OF
ADJUSTMENT BASED ON INTERVIEWS

Variables	Age	Boys	Girls
Total Adjustment	11	+.27	-.08
Signs With Inter-	13	+.36*	-.05
view "Self" plus	15	-.14	+.08
"Social" Adjustment	18	-.29	+.03
Davidson "Self"	11	+.04	+.02
Adjustment With	13	+.28	-.25
Interview "Self"	15	-.22	-.05
Adjustment	18	-.49*	-.18
Davidson "Social"	11	+.16	-.13
Adjustment With	13	+.23	+.06
Interview "Social"	15	+.01	+.24
Adjustment	18	+.04	+.08

*indicates correlations significant at the .01 level of confidence.

The correlations of total adjustment scores and interview ratings demonstrate no relationship between the level of adjustment indicated by the Rorschach signs and the case workers' evaluations of the adjustment of children with whom they have worked. In general, they vary non-significantly around zero, as do the correlations between Rorschach signs and interview ratings of self and social adjustment. Of the 24 correlations, only two are significantly different from zero (those doubly underlined in the table).

¹⁵The descriptions and definitions of the codes which governed the case workers' ratings are presented in the appendix. For the correlations with the Davidson scale, the ratings for social adjustment *A* and *B* were combined. The interviews upon which the ratings were based are described in section *C*.

Insofar as we can tell from her monograph (6), Davidson's division of the items into the two adjustment scales was based not on empirical findings but on personal assumption. In some cases, the resulting placement of items is questionable. For example, one could argue that the sign $\text{Sum } C > Fc + c + C'$ refers as much to social adjustment as to personal adjustment, where Davidson has placed it. The same is true of the sign $P > 3$.

For these reasons, the authors feel that the most relevant data regarding the relationship between the Davidson Adjustment Signs and our external criteria of adjustment is to be found in the first group of correlations, those between the total adjustment sign score and the combined interview ratings. On this basis, the obvious conclusion is that no relationship is demonstrated; that the number of Davidson signs in a subject's Rorschach does not predict the degree of adjustment revealed in personal interviews.

Table 12 summarizes the differences in quartile placement on the Rorschach adjustment signs and the interview ratings and the direction of those differ-

TABLE 12
DISCREPANCY BETWEEN QUARTILES ON RORSCHACH ADJUSTMENT SIGNS AND TOTAL OF
INTERVIEW RATINGS AVERAGED FOR ALL AGES

	Boys N = 49	Girls N = 48	Total N = 97
Difference of 2 or 3 quartiles	41%	44%	42%
Rorschach quartile higher	24%	23%	24%
Interview quartile higher	16%	21%	19%
Difference of 1 quartile	35%	33%	34%
Rorschach quartile higher	16%	17%	16%
Interview quartile higher	18%	17%	18%
No difference	24%	23%	24%

ences. The Rorschach quartile and the interview quartile were the same in only one-fourth of the cases. In over 40 per cent of the cases, there was a difference of two or three quartiles, with the Rorschach quartile higher in about half of these and the interview quartile higher in the other half. These findings illustrate from another point of view the lack of any demonstrable relationship between the adjustment signs and the interview ratings.

D. SUMMARY AND CONCLUSIONS

We have presented data on the incidence of Rorschach "neurotic" and "adjustment" signs in a population of normal adolescents who were tested

at ages 11, 13, 15, and 18. Age changes and sex differences in sign incidence have been presented and discussed.

The incidence of many of the signs currently used to diagnose neurosis was found to be considerably higher than would be expected if the signs were truly indicative of neurosis. More than one-third of a group of 145 18-year-olds would be diagnosed as neurotic on the basis of having five or more signs. Three of the 14 signs used showed significant decrease of incidence with age. Age changes over the span of 11 to 18 years of the other 11 signs did not permit the assumption that the incidence was a function of adolescent immaturity and would decrease with increasing age. The girls showed fewer signs than the boys over the entire age range, and the magnitude of the differences even at age 18 suggests that sex differences on the Rorschach deserve more attention than they have received. On the whole, the characteristics of these "neurotic" signs in a normal group make questionable their validity and usefulness as a diagnostic technique with adolescents, and indicate a need for further study of normal adults.

The 17 "adjustment" signs vary considerably in terms of incidence. Some show a uniformly high incidence while others never occur in more than one-third of any age or sex group. As a whole, the signs do not occur with the frequency to be anticipated in a group of well functioning adolescents. There is a consistent sex difference in the proportion showing the adjustment signs, the difference being always in favor of the girls. A few of the signs show an increase of incidence with increasing age, indicating a probable relationship to maturity. The signs which show statistically significant age changes and sex differences are primarily those concerned with the variables of movement, color, and number of responses.

The correlations between scores on the adjustment sign list and interview ratings of adjustment are largely non-significant and indicate that no relationship exists between these Rorschach signs of adjustment and the adjustment of the child as evaluated by case workers.

We must conclude that the sign lists do not hold up as indicators of adjustment or the lack of it in our random sample of adolescents, and offer the following hypotheses as to why they do not.

The first is in terms of a common criticism of "sign" methods as applied to the Rorschach. Sign methods are generally attempts to extract Rorschach variables from their total setting and assign them quantitative values. Many sophisticated Rorschach workers contend that this procedure destroys the significance of the test variables and that, for this reason, sign methods will

not be productive in work with either individuals or groups. Our study of two "sign" lists leads us to favor this hypothesis.¹⁶

A second consideration is the inadequacy of previously available Rorschach data on well selected and thoroughly studied normal groups. We have far more data on the Rorschachs of disturbed subjects and are perhaps inclined to arrive at "norms" by a process of eliminating those relationships and responses found in clinical cases. It is possible that the items in the Davidson list (which were not validated by her against an external criterion of adjustment) represent "ideal" rather than "average" Rorschachs and, therefore, will not occur with any great frequency in a normal group.

In the same vein and apropos of the neurotic signs, the question has been

TABLE 13
PER CENT OCCURRENCE OF RORSCHACH NEUROTIC SIGNS
SERIES B

Sign		Age in years		
		12	14	16
(1) $R < 25$	b	81	88	81
	g	89	94	76
(2) $M < 2$	b	44	36	29
	g	40	44	44
(3) $FM > M$	b	69	52	14
	g	77	56	44
(4) Color shock	b	72	72	62
	g	64	44	32
(5) Shading shock	b	31	20	19
	g	19	21	28
(6) Refusals	b	25	28	33
	g	30	21	16
(7) $F\% > 50$	b	14	32	48
	g	6	18	56
(8) $A\% > 50$	b	53	48	48
	g	60	74	76
(9) $FC < 2$	b	75	84	86
	g	70	76	84
(10) No FC	b	58	68	67
	g	47	50	52
(11) $R < 12$	b	33	28	43
	g	43	35	32
(12) $FM \leq 2M$	b	47	28	14
	g	51	41	32
(13) No FM	b	3	12	19
	g	2	0	8
(14) $A + At\% > 65$	b	22	28	24
	g	34	32	40

¹⁶The next paper in this series will present a study of subjective ratings of adjustment by Rorschach experts with training and experience in evaluating the Rorschach Gestalt.

raised as to whether the Rorschach (or the clinician who works with it) is not biased in the direction of abnormality. It is often evident in a clinical study that, while the experimental group is selected on the basis of known characteristics, the group designated as a control group consists of individuals who are *not* known to be deviant rather than known to be "normal." Other studies simply contrast the Rorschachs of two or more deviant groups without reference to the content of Rorschachs in a normal group. Do we, then, see too readily the signs of disturbance while not recognizing the contexts of equilibrium?

TABLE 14
PER CENT OCCURRENCE OF RORSCHACH ADJUSTMENT SIGNS
SERIES B

Sign		Age 12	Age 14	Age 16
		Total group N = 83	Total group N = 60	Total group N = 46
(1) $R > 20$	b	28	16	24
	g	15	15	28
(2) No rejections	b	75	72	67
	g	70	79	84
(3) $M \leq FM$	b	28	44	76
	g	23	44	56
(4) $M > 2$	b	50	44	57
	g	43	50	32
(5) $FK + Fc > 1$	b	39	24	43
	g	40	21	36
(6) $FC \leq CF$	b	33	20	29
	g	45	44	44
(7) $FC > 1$	b	25	16	14
	g	29	24	16
(8) No C	b	97	96	90
	g	98	100	92
(9) Sum $C > Fc + c + C'$	b	33	16	29
	g	29	21	8
(10) $P > 3$ ($< 30\%R$)	b	33	24	19
	g	17	12	28
(11) $F\% \leq 50$	b	86	68	52
	g	94	85	44
(12) $A\% \leq 50$	b	47	52	52
	g	40	26	24
(13) %VIII-X is 40-60	b	33	48	14
	g	38	32	40
(14) $W:M = 2:1$	b	31	12	10
	g	17	18	20
(15) $Dd + S\% \leq 10$	b	75	84	86
	g	85	91	76
(16) No color shock	b	28	28	38
	g	35	56	68
(17) No shading shock	b	69	80	81
	g	81	79	72

TABLE 15
APPENDIX

Adjustment to self (size, sex, physical make-up, abilities, disposition, past misdemeanors, etc.)

- (1). Extreme lack of confidence shown either through excessive self-depreciation, unfavorable comparison, or marked braggadocio. Intense guilt feelings, or feelings of inadequacy
- (2). Not as marked as 1 but marked unacceptance of one or two items of make-up. However, if to the point of dominating whole personality, classify under 1. Or many mild points of discomfort. Over-adjusts to others at own expense.
- (3). Lacks confidence in some particulars but for the most part is unselfconscious or unaware of assets or liabilities. Occasional tension or conflict about self but not characteristic.
- (4). Recognizes handicaps but accepts them realistically and without much tension. Little evidence of compensatory behavior.
- (5). Full of confidence — accepts limitations and assets easily.

Social adjustment

A. (Refers, as used here, to social techniques, ease with which a child evokes acceptance, approval, friendliness or loyalty; or incurs disapproval, distrust or hostility. Emphasis is on a social criterion rather than on child's security in social situations — this last being handled under such items as shyness, friendliness, quarrelsomeness, etc.)

- (1). Extremely unsuccessful in social relations. Intensely disliked or resented by practically all of his playmates. Very poor social techniques and constantly does things which offend his social group. Actively avoided.
- (2). Unsuccessful in social relationships; not popular but does not give rise to the intense reactions of 1. Ignored or not spontaneously included in group play, etc. Does things which annoy. Definitely disliked by many children; has some supporters.
- (3). Not outstanding. Normally successful in having people like him. Some playmates may not like him but has approving supporters. Or successful with one sex but unsuccessful with the other.
- (4). More popular than the average; has good social techniques.
- (5). Unusually successful in social relations. Inspires friendliness or loyalty. Has many people devoted to him. Constantly gives rise to such statements as, "Isn't he a swell guy?" "She's a grand person," etc.

B. (Refers, as used here, to child's feeling of tension or confidence and satisfaction in social relations)

- (1). Acutely uncomfortable, dislikes people, feels he is treated unfairly or is disliked by others. Chronically sulks or is aloof.
 - (2). Definite but less acute tension than 1. Feels he is not liked as well as other children. "Children don't play with me much." Or worried because one or two children dislike him. Insecurity may take form of shyness or aggressive attempts at social overtures.
 - (3). Takes his social relationships for granted, doesn't occur to him to consider whether he is liked or disliked. Some frictions but "have as many friends as most children do." May have few friends but apparently has no tension or worry on this score. Feels secure with one sex but a bit uneasy with the other.
 - (4). More confidence in social relationships than average. Very little tension and many satisfactions.
 - (5). Likes children very much; enjoys being with them. Is enthusiastic and happy in relations. Confident, feels socially secure.
-

A third factor undoubtedly involved in our results with the adjustment signs is the complexity of the concept of adjustment. Our external criteria are ratings of adjustment based on behavioral data, observations, and inferences over a wide range of situations and circumstances. Does overall adjustment in this framework mean the same thing as adjustment as revealed by the Rorschach? An attempt to answer this question would lead the authors into discussions beyond the scope of this paper. It is worth considering, however, how much agreement can be expected between any two measures of such a generalized concept.

A fourth hypothesis is that the Rorschach is, in fact, not valid in the sense of predicting degree of adjustment. Our evidence is by no means conclusive even in this limited area. For the present, the authors retain their faith in the Rorschach as a sample of significant behavior, but reiterate their impression that many stereotyped approaches to it should be scuttled.

REFERENCES

1. ASH, P. The reliability of psychiatric diagnoses. *J. Abn. & Soc. Psychol.*, 1949, **44**, 272-276.
2. BECK, S. J. Personality structure in schizophrenia: A Rorschach investigation in 81 patients and 64 controls. *Nerv. Ment. Dis. Monogr. Ser.*, 1938, **63**, ix+88.
3. BELL, J. E. Projective Techniques. New York: Longmans, Green, 1948.
4. BROSN, H. W., & FROMM, E. O. Rorschach and colorblindness. *Rorschach Res. Exch.*, 1940, **4**, 39-70.
5. DAVIDSON, H. H. Personality and Economic Background: A Study of Highly Intelligent Children. New York: King's Crown Press, 1943.
6. ———. A measure of adjustment obtained from the Rorschach protocol. *J. Projec. Tech.*, 1950, **14**, 31-38.
7. EDGERTON, H. A., & PATTERSON, D. G. Table of standard errors and probable errors of percentages for varying numbers of cases. *J. Appl. Psychol.*, 1926, **10**, 378-391.
8. GRANT, M. Q., RANZONI, J. H., & IVES, V. Reliability and validity of judges' ratings of adjustment on the Rorschach. (To be published.)
9. GUIRDHAM, A. The Rorschach test in epileptics. *J. Ment. Sci.*, 1935, **81**, 870-893.
10. HARROWER-ERICKSON, M. R. The value and limitations of the so-called "neurotic signs." *Rorschach. Res. Exch.*, 1942, **6**, 106-115.
11. HERTZ, R. Personality patterns in adolescence as portrayed by the Rorschach ink-blot test method: II. The color factors. *J. Gen. Psychol.*, 1943, **28**, 3-61.
12. HERTZMAN, M., & MARGULIES, H. Development changes as reflected in Rorschach test responses. *J. Genet. Psychol.*, 1943, **62**, 189-215.
13. KLOPPER, B., & KELLEY, D. M. The Rorschach Technique. Yonkers: World Book, 1942.
14. MACFARLANE, J. W. Studies in child guidance: I. Methodology of data collection and organization. *Monog. Soc. Res. Child Devel.*, 1938, **3**, No. 6, Pp. 254.
15. MCFATE, M. Q., & ORR, G. Through adolescence with the Rorschach. *Rorschach Res. Exch. & J. Projec. Tech.*, 1949, **13**, 302-319.

16. MCNEMAR, Q. *Psychological Statistics*. New York: Wiley, 1949.
17. MIALE, F. R., & HARROWER-ERICKSON, M. R. Personality structure in the psychoneuroses. *Rorschach Res. Exch.*, 1940, **4**, 71-74.
18. MUNROE, R. Inspection technique. *Rorschach Res. Exch.*, 1941, **5**, 166-191.
19. PIOTROWSKI, Z. A. Positive and negative Rorschach organic reactions. *Rorschach Res. Exch.*, 1940, **4**, 147-151.
20. RANZONI, J. H., GRANT, M. Q., & IVES, V. Rorschach "card-pull" in a normal adolescent population. *J. Projec. Tech.*, 1950, **14**, 107-133.
21. ROSS, W. D. A contribution to the use of "signs" in the Rorschach method. (Paper read at second annual meeting of the Rorschach Institute, 1941.)

Institute of Child Welfare
University of California
Berkeley 4, California

VISUAL DISCRIMINATION PERFORMANCE IN RHESUS
MONKEYS FOLLOWING EXTIRPATION OF PRESTRIATE
AND TEMPORAL CORTEX*^{1,2}

Departments of Psychology and Anatomy, Emory University

ARTHUR J. RIOPELLE AND HARLOW W. ADES

A. INTRODUCTION

The cortical regions immediately rostral to the visual projection area of primates are referred to as the para- and peri-striate areas (18 and 19 of Brodman). Clinical observations have led to the interpretation that this is a "visual association" area. Supporting evidence is found in the cortical neuronographic studies of Bonin, Garol, and McCulloch (3). The effects of extirpations of this region in monkeys have been reported by Ades (1), Ades and Raab (2), and by Lashley (10).

Ades (1) trained three rhesus monkeys to discriminate between pairs of visual stimuli on the basis of size, shape, and color. He summarized his results as follows: "Monkeys so trained lose the discrimination habit if areas 18 and 19 are destroyed bilaterally in one stage. The discrimination can then be relearned at approximately the same rate as originally. If the cortical destruction is carried out in two stages with an interval between (18 days) during which testing is continued, the habit is lost after neither the first nor the second operation."

Corroborative data have been presented by Ades and Raab (2) using *F* versus inverted *F* as test stimuli. Approximately twice as much postoperative as preoperative training was required to attain equal levels of performance in the bilateral animals. Again, with training trials presented between operations, seriatim destruction of areas 18 and 19 produced no deficit in performance after either operation. Subsequent bitemporal decortication resulted in a postoperative amnesia which could not be overcome with prolonged training. Both learning and retention are good following bitemporal decortication, although sub-

*Received in the Editorial Office on June 11, 1951.

¹Supported in part by the Research Committee of the Graduate School of the University of Wisconsin from funds supplied by the Wisconsin Alumni Research Foundation, and by the National Institutes of Health.

²This experiment was carried out at the University of Wisconsin primate laboratory.

sequent bilateral removal of areas 18 and 19 permanently abolished the form discrimination.

Lashley (10), however, reported that spider monkeys did not lose the ability to discriminate either colors or forms following bilateral extirpation of areas 18 and 19. Three monkeys trained to choose between two stimulus-objects, a green square and a green circle, required a total of 151 preoperative trials to learn the discrimination, but only 40 trials were needed to re-acquire the habit postoperatively to the same criterion of 20 consecutive errorless trials. He suggested three possible explanations for the differences between his and Ades' results: (a) differences in genera (Ades: *Macaca*; Lashley: *Ateles*), (b) differences in lesions, and (c) differences in motivation (Lashley: punishment for errors; Ades: reward for correct responses). Lashley believed that the last is probably the most important difference but has since stated in a personal communication that differences in recuperation period might also be important.

Two other experimental differences might be noted in attempting to understand the inconsistency of results. The first is the variation in difficulty or nature of the tasks. Ades (1) and Ades and Raab (2) used patterns, whereas Lashley used stimulus-objects. These differences could well be important. Indeed, differences in difficulty between kinds of pattern stimuli might be important. Thus Ades' animals relearned dissimilar patterns at the same rate as originally, but Ades and Raab's animals relearned at a slower rate when mirror-image patterns were used. Similar differences can be noted for the original learning of the two dissimilar problems. Harlow (6) had previously shown that for normal rhesus monkeys, pattern discrimination is more difficult than object discrimination and that mirror-image patterns are learned less readily than dissimilar patterns.

There is evidence that the extent of loss is related to the difficulty of the task. The retarded relearning which occurred in the Ades and Raab study points in this direction. Lashley, too, has noted that some of his monkeys postoperatively performed with less efficiency on the more difficult visual tasks.

The second difference is the unequal amount of laboratory discrimination testing experience of the animals in the three studies reported. Lashley's animals had extensive discrimination training prior to the experiment in question whereas the only laboratory experience of Ades' animals was the specific pretraining for that particular experiment. There is reason to believe that both amount of preoperative training and level of preoperative per-

formance are related to the extent of loss following cortical assault. Harlow (7) has shown that efficient "learning sets" may be developed following extensive cortical lesions, and unpublished studies from his laboratory clearly show that previous training greatly influences the performances of animals subsequently subjected to cortical lesions.

B. PURPOSE OF THE PRESENT STUDY

The purpose of the present study was to extend the work of the previous investigators in the following ways: (a) Increase the range of visually guided behaviors measured since the sensitivity of a measure of cortical damage is presumably related to the difficulty of the task. (b) Assess the rôle of operation on learning as well as retention of visual discrimination problems. (c) Determine the effects of simultaneous bilateral extirpation of the preoccipital area and the temporal lobes, since the seriatim effects of these operations had been noted by Ades and Raab (2). Furthermore, Klüver and Bucy (8, 9) found that bilateral extirpation of the temporal lobes of monkeys resulted in what they termed "psychic blindness." Their animals showed an apparent inability to recognize objects visually.

C. TESTS AND TEST STIMULI

Five animals were trained to discriminate between pairs of visual stimuli of the following six types: stimulus-objects differing in form (Test I), patterns differing in form (Test II), mirror-image stimulus-objects (Test III), mirror-image patterns (Test IV), identical form objects differing in color (Test V), and a series of stimulus-objects (Test VI). The stimuli used in this experiment are illustrated in the graphs. The stimuli for the dissimilar stimulus-object discriminations (Test I) were two different shapes cut out of unpainted masonite. The dissimilar patterns (Test II) consisted of different white patterns painted on each of two black wedges. The mirror-image stimulus-objects (Test III) consisted of a black wooden letter *K* standing upright upon a square white base and a reverse *K* on the other white base. The mirror-image patterns (Test IV) were block letters of the alphabet (*C*, *F*, or *L*) painted black on one white wedge and their mirror images painted on the other.

The dissimilar colors (Test V) were made up of 12 three-inch square blocks of wood $\frac{3}{8}$ inches in height. Each was painted one of three different brightnesses of red, green, blue, or gray. In this test the red block was always rewarded. On the first day the red blocks were paired against the

gray; on the following day each red block was paired against each of the three shades of the other two hues. All combinations of brightnesses were used in a balanced order to preclude the possibility of discrimination on the basis of cues other than hue.

The series of different objects (Test VI) was made up of 10 pairs of stimuli taken from the laboratory's stock of 500 pairs of randomly arranged stimulus-objects varying in multiple dimensions. One object of each pair was arbitrarily chosen as correct and its selection was rewarded. Each pair of stimuli constituted a problem which was run for only six trials.

D. APPARATUS AND GENERAL TEST PROCEDURE

The general test situation was the same as that described in previous studies carried out at the University of Wisconsin (4, 5). The animal is placed in a special testing cage which fronts a table on which is placed a sliding tray bearing the test objects. Two screens are used. One, the forward opaque screen, is lowered immediately in front of the animal's cage while the food wells are baited. The other, a one-way vision screen, is located in front of the experimenter. It is lowered during the time the animal is making a choice.

The sliding tray was light gray in color and had two food wells set 15 inches apart, center to center. The food wells were sufficiently shallow so that a small piece of food in an uncovered well was readily visible from the monkey's position, yet deep enough to allow placement of the test object over the baited well without its being elevated.

By appropriate use of screens the animal could not see the food wells being baited nor the experimenter while making a choice.

E. DESIGN OF EXPERIMENT

The experiment was carried out in two parts. In Part I, three animals (105, 106, 108) were operated prior to complete learning of the visual discriminations. The other two (101, 107) served as controls. In Part II, Monkeys 101 and 107 were then operated after complete learning and subsequently retrained. The prior test history of these animals has been reported (7).

F. ANIMALS AND OPERATIONS

Five mature rhesus monkeys, well tamed and with at least four years of continual test experience involving hundreds of discrimination problems, (7) served as subjects.

Extirpation of the cortex was carried out aseptically under nembutal anesthesia. The cortex was removed by suction through a small glass pipette. Great care was taken to remove only gray matter, which, in a blood free field, is readily distinguished from white matter. The pairings of animals and operations are indicated in Table 1. This table also shows the date of operation and the number of recuperation days allowed for each monkey prior to postoperative testing.

TABLE 1

Animal	Date of operation	Extent of lesion	Recuperation period
106	7/8/48	Bilateral 18, 19	7
108	7/8/48	Unilateral 18, 19	7
108	8/3/48	*Bilateral 18, 19	2
105	7/8/48	Bilateral 18, 19 and bilateral temporal	15
101	3/10/49	Bilateral 18, 19	3
107	3/7/49	Unilateral 18, 19	3
107	3/21/49	*Bilateral 18, 19	5

*Denotes extent of total lesion.

The operations of Animals 105, 106, and 108, with respect to areas 18 and 19, invaded the medial surface of the brain as far as was easily accessible. This distance was not very great because of the configuration of the cortex at the convergence of the lunate and postparietal sulci with the deep pial banks. Inferiorly it was possible to reach as far as the point where the inferior occipital sulcus passes out of sight on the ventral surface of the occipital lobe. According to the cytoarchitectonic maps the concealed medial and ventral portions of areas 18 and 19 escaped injury.

The operations of Animals 101 and 107 did not invade the medial surface of the hemisphere.

G. RESULTS

The results for Part I in which the operations were carried out during the course of learning are first considered for successive tests.

1. *Dissimilar Stimulus-Objects (Test I)*

In Figure 1 are presented the per cent correct choices for successive blocks of 50 trials each. The preoperative training on this test for the experimental animals amounted to 75 trials. The three operated animals exhibited consistently poorer performance than the controls. In addition, the extent of loss appears related to the magnitude of the lesion. The second operation of Animal 108 produced no loss in performance on this test.

This test is relatively easy as can be seen from the two control curves, both of which start at a high value and rapidly approach 100 per cent correct.

2. Dissimilar Patterns (Test II)

Preoperative training on this test amounted to 75 trials. The test results are shown graphically in Figure 2.

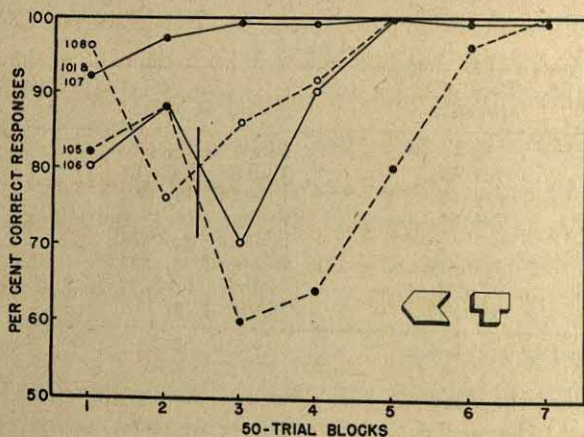


FIGURE 1

DISSIMILAR STIMULUS-OBJECTS

Animals 105, 106, and 108 were operated after Test Period 2. Animal 108 was operated again after Test Period 5. Animals 101 and 107 served as normal controls.

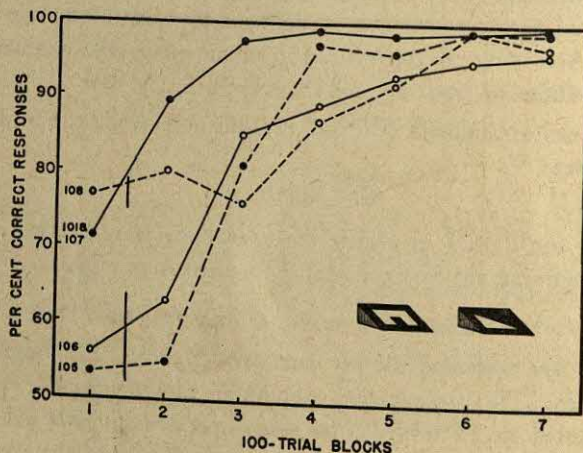


FIGURE 2

DISSIMILAR PATTERNS

Animals 105, 106, and 108 were operated after Test Period 1. Animal 108 was operated again after Test Period 3. Animals 101 and 107 served as normal controls.

The postoperative performance of the operates is inferior to that of the controls. There appears to be an immediate difference between the unilateral animal and the bilateral animal in favor of the lesser lesion, but this difference is quickly dissipated. After the first 100 postoperative trials differentiation among the operates on the basis of test performance is impossible.

In this test, as in the dissimilar objects test, removal of the contralateral areas 18 and 19 resulted in no measurable loss in performance of Animal 108.

It will be noted that this test is more difficult than the dissimilar stimulus-object test; all animals requiring a greater number of trials to attain a given level of performance.

3. *Mirror-Image Objects (Test III)*

This test was not introduced into the battery until after the experimental animals were operated. As a result, preoperative data are not available. In fact, Animal 108 had undergone a second operation before it had experience with this test.

Test performances for each animal are graphically indicated in Figure 3. No systematic differences appear between controls and operates or among the operates.

This test appears to be of intermediate difficulty; requiring more trials to learn than Tests I and II yet fewer than Test IV.

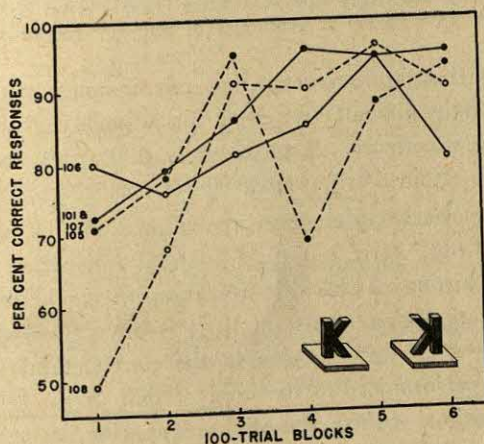


FIGURE 3
MIRROR IMAGE STIMULUS-OBJECTS

All of the animals except Control Animals 101 and 107 were operated (108 twice) prior to testing.

4. *Mirror-Image Patterns (Test IV)*

Three tests of this type were actually employed. In one, the monkeys were required to select the wedge on which was painted an *L* and to reject the wedge on which was painted a lateral mirror-image *L*. In the second test, a *C* was paired against its mirror image, and on the third, an *F* against its inverted mirror-image. The results of all three tests were highly similar and for this reason the data for the *L* test only are shown in Figure 4.

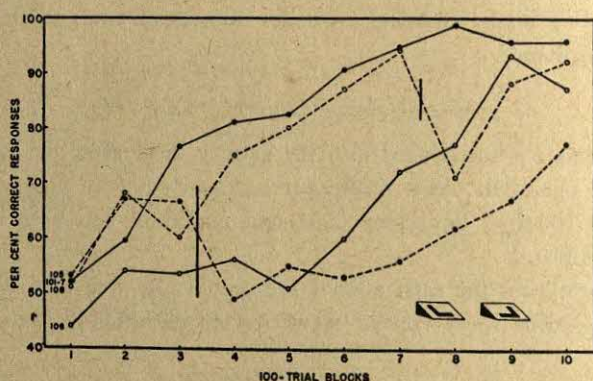


FIGURE 4
MIRROR IMAGE PATTERNS

Animals 105, 106, and 108 were operated after Test Period 3. Animal 108 was operated again after Test Period 7. Animals 101 and 107 served as normal controls.

Examination of this figure reveals that mirror-image pattern discrimination is an exceedingly difficult test even for animals having extensive discrimination testing experience. Approximately 600 trials were required for the two controls to attain a criterion of 90 per cent correct responses.

These tests most clearly differentiate the animals in this experiment. Four facts seem to stand out. First, one-stage bilateral extirpation of these regions seriously retards learning of this discrimination. Second, unilateral extirpation of the prestriate region results in little if any retardation in learning. Third, extirpation of the same region in the contralateral hemisphere at a later date reduces performance to the level of that of the bilateral one-stage operate. In this animal (106) the slight depression in performance prior to operation cannot be so considered. Fourth, simultaneous one-stage bilateral extirpation of both areas 18 and 19 and the temporal lobes retards learning to a still greater degree.

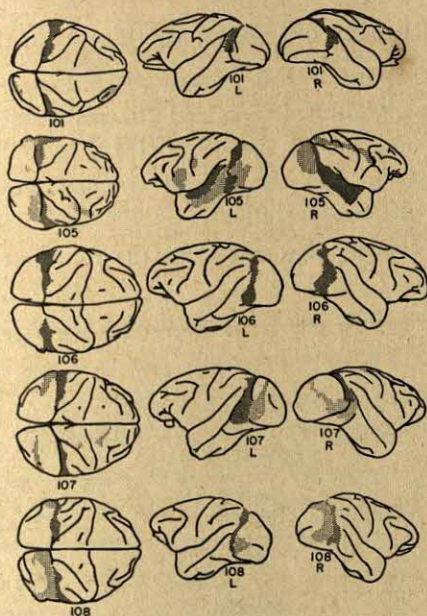


FIGURE 5

EXTENT OF LESIONS

These drawings represent tracings from photographs of the cerebral hemispheres after perfusion and fixation in formalin. The hatched areas indicate deep lesions made by pipetting cortex under suction. The stippled areas indicate inadvertent superficial lesions, such as dural adhesions, and the like. None of these are deep lesions, although in one or two cases there is considerable question as to the functional capacity of such areas because they were actually herniated through defects in the dura.

Monkey No. 101. Has bilateral lesions of the preoccipital cortex, principally in area 19, with encroachment on area 18. The infero-lateral, inferior, and medial parts of 18 and 19 are intact. The posterior lip of the lunate sulcus was left intact over most of its extent on each side except that the distal third on the right side was destroyed and a small area proximally on the left was likewise destroyed. Otherwise, the lesions were confined to the superficial aspects of area 19.

Monkey No. 105. Bilateral lesions of the preoccipital cortex, much more extensive than in animal No. 101, although the inferior portions of areas 18 and 19 and the medial portions are intact. The posterior lip of the lunate sulcus is destroyed on both sides and dural adhesions make some encroachment on area 17 on both sides. Temporal lesions are entirely confined to the superior temporal gyrus, which is largely removed on both sides.

Monkey No. 106. Bilateral lesions of the preoccipital cortex, including the posterior lip of the lunate sulcus on both sides, over all of its exposed extent. These are the most extensive lesions of any of the group. Again, the medial parts of 18 and 19 and a little of inferior portions remain.

Monkey No. 107. Bilateral lesions of areas 18 and 19 with the posterior lip of the lunate sulcus intact. There is very slight encroachment on the posterior lip only on the left side distally. The lesion on the left is larger than that on the right, but large parts of areas 18 and 19 remain undamaged.

Monkey No. 108. Complete destruction of the superficial portions of areas 18 and 19 as well as the entire posterior lip of the lunate sulcus on each side. Inferior and medial parts of 18 and 19 are not damaged.

H. RESULTS ON POSTOPERATIVE RETENTION

After Animals 101 and 107 had learned the discriminations employed in this study they too were operated.³ Animal 101 was subjected to a one-stage bilateral extirpation of areas 18 and 19, and 107 to a two-stage operation involving similar regions. For this animal training was continued during the interval between the two operations.

Preoperative and postoperative scores for the four tests mentioned above are presented in Table 2.

From this table it can be seen that transoperative retention for all tests is consistently high.

TABLE 2
NUMBER OF CORRECT DISCRIMINATIONS IN LAST 100 PREOPERATIVE AND FIRST 100 POST-
OPERATIVE TRIALS

Test	Animal	Extent of total lesion	Preoperative 100 trials	Postoperative 100 trials
I. Dissimilar Objects	101	B* 18, 19	100	98
	107	U 18, 19	100	99
	107	B 18, 19		100
II. Dissimilar Patterns	101	B 18, 19	100	98
	107	U 18, 19	97	97
	107	B 18, 19		99
III. Mirror Image Objects	101	B 18, 19	96	90
	107	U 18, 19	99	93
	107	B 18, 19		89
IV. Mirror Image Patterns	101	B 18, 19	98	96
	107	U 18, 19	98	89
	107	B 18, 19		90

*B denotes bilateral, U denotes unilateral, and refers to extent of total lesion.

1. Dissimilar Colors (Test V)

Pre- and post-operative results for the color discrimination test are summarized in Table 3. Retention data for Animal 107 (two stages 18, 19) was not obtained until after the second operation.

These results give no indication of loss occurring in color discriminations following extirpation of the prestriate cortex alone, either unilateral or bilateral, or in conjunction with the temporal lobes. In fact, more gains than losses occurred and there is no correlation between lesion size and differential score.

³We are indebted to Dr. Paul Settlage of the University of Wisconsin for performing these operations.

TABLE 3
NUMBER OF CORRECT COLOR RESPONSES IN 54 TRIALS BEFORE AND AFTER OPERATION

Animal	Preoperative scores		Postoperative scores	
	Red vs. Black	Red vs. Blue or Green	Red vs. Black	Red vs. Blue or Green
108	43	45	53	54
108			54	54
101	51	54	53	47
106	54	54	53	54
107	49	50	46	53
105	47	49	47	54

2. Multidimensionally Different Objects (Test VI)

It was indicated earlier that the multidimensionally different objects test consisted of a series of six-trial problems. All animals had previously been trained extensively on similar problems. After this training the animals required only one trial to learn to select the correct object. Theoretically perfect performance for 60 trials (10 problems) would be 55 correct choices. Examination of Table 4, which presents the data on this test, indicates that postoperative performance was nearly perfect.

Animal 108 lost ground after the first operation but regained after the second. Animal 107 lost little or nothing after either of the operations. Of the bilateral one-stage animals, one (106) possibly lost slightly, the other (101) probably did not. Animal 105 showed some decrement in performance following operation but even this is slight. Whatever loss did occur was well within the range of normal fluctuations of these animals.

I. DISCUSSION

The principal result of this investigation is that bilateral extirpation of areas 18 and 19 alone and in conjunction with extirpation of portions of

TABLE 4
NUMBER OF CORRECT PRE- AND POST-OPERATIVE RESPONSES TO STIMULUS OBJECTS VARYING IN MULTIPLE DIMENSIONS FOR 10 PROBLEMS (60 TRIALS)

Animal	Total lesion	Preoperative	Postoperative
101	B* 18, 19	54	51
105	B 18, 19 + temporal	51	42
106	B 18, 19	52	46
107	U 18, 19	55	54
107	B 18, 19		51
108	U 18, 19	54	46
108	B 18, 19		53

*B denotes bilateral, U denotes unilateral and refers to extent of total lesion.

temporal lobes in the rhesus monkey seriously retards the learning of mirror-image pattern, dissimilar object, and dissimilar pattern discriminations. The learning of mirror-image object discriminations appears not to be retarded following any of the above operations.

Unilateral extirpation of the prestriate region yields slight, if any, loss in the learning of dissimilar or mirror-image pattern and object discriminations. Whereas the first unilateral extirpation in Animal 108 resulted in no decrement in the learning of mirror-image pattern discriminations, the second operation performed on the opposite hemisphere reduced performance to the level of that of the one-stage bilateral animal (106). Subsequent learning was rapid.

Unilateral and bilateral extirpations of the prestriate regions *after* learning to a high criterion resulted in little or no loss on any of the tests.

Certain similarities and differences of results between the present and previous investigations deserve to be noted. First, the previous investigators (1, 2, 10) operated on their animals after the animals had learned the discrimination problems. The present experiment, by focusing attention upon postoperative learning, may be considered an extension of the previous work. The differences obtained in the present study in postoperative learning support Ades' (1) interpretation that areas 18 and 19 and the temporal lobes are "essentially involved" in the development of the particular patterns of visual interpretation or association employed in these experiments. The fact that retarded learning did not occur for all tests indicates that a general motivational loss did not occur. Such an interpretation is supported by observations of the animals' behavior in the testing situation. In all cases the operated animals, when ready for testing, behaved in a manner entirely appropriate to the test situation. They appeared alert and responded quickly to the test objects.

The decrease in performance on the part of Animal 108 following the second unilateral extirpation has not previously been noted. It occurred in only the mirror-image pattern test. This decrement, reducing performance to the level of a one-stage bilateral animal, cannot be considered an artifact or chance result, for the identical event took place on *all three* of the mirror-image pattern discrimination tests. It is to be noted that the second operation for this animal occurred before complete learning had taken place.

After prolonged near-perfect performance, the effect of extirpation of the prestriate regions is quite different. Unilateral and bilateral removal of those areas resulted in little or no loss on any of the tests. The unilateral

results on retention agree with those of Ades (1) and Ades and Raab (2). The bilateral results are in striking disagreement with the data of these workers and do agree with those of Lashley.

Two important facts must be kept in mind when considering these differences. First, the volume of tissue removed from the animals on whom retention data are available was definitely less than that removed from the animals used in the learning study. It will be remembered that for Animals 101 and 107, among other things, the medial surfaces were not touched. Second, the animals in the Ades (1) and Ades and Raab (2) studies had little or no prior testing experience, whereas the animals in the present study and in that of Lashley (10) had extensive discrimination training prior to the investigations under consideration.

The sophisticated animals of the present study had a tremendous amount of visual discrimination testing experience, some of which, undoubtedly, can be related to these particular discrimination problems. In addition, they had the specific training given during the acquisition period. Animals such as 101 and 107 and the spider monkeys of Lashley when operated had extensive reserves of both the general and the specific training. Animals 105, 106, and 108, on the other hand, had little of the specific training and much of the related training. At the other extreme is the group of animals used in Ades' laboratory. These animals had extensive specific training but little generalized training. When viewed in this light, the results from the different laboratories appear to fall in line. Apparently experience is an important variable. The effects of specific training for a single discrimination does not survive cortical assault. Generalized training permits the animal to learn a particular problem at moderately reduced efficiency. Extensive reserves of both kinds of training result in highly efficient postoperative performance. Such an hypothesis accounts for Lashley's failure to obtain loss, and also the failure to get loss of retention in the present investigation. It also accounts for the definite but intermediate loss in learning of the present study and the great loss obtained by Ades and Ades and Raab.

In stressing this point, however, one should not lose sight of the differences in performance between unilateral and bilateral operations. In every case where loss occurred, that loss was greater following bilateral extirpation than following unilateral extirpation. In addition, the greater deficit of Animal 105 (bilateral 18, 19 + temporal) clearly establishes the importance of these regions, especially through their interactions, for the learning of visual discriminations.

One of the hypotheses of the present study was that sensitivity to cortical damage might be related to the difficulty of the task. In general, this hypothesis seems to be supported, although there is one notable exception. The mirror-image objects test, which, on the basis of the number of trials to criterion, appeared to be of intermediate difficulty, and hence of some predictive value, failed to differentiate between the operated and controls and among the operated. Such a finding suggests that these tests have a diagnostic value, and may be useful in determining the *kinds* of visual functions for which learning is most retarded following cortical assault.

In general the results of the present study, when taken in combination with those of previous studies, especially as to the differences which exist between learning and retention following cortical injury, amplifies the implications for human operations of a similar nature. Although the individual animal may not suffer greatly on specific visual discrimination problems, especially if he has a wealth of similar experience behind him, he is, nevertheless, less able to accommodate himself to a new visual discrimination problem. The first expectation would be that the results would be in a similar direction and, most probably, would be even more striking in human subjects.

J. SUMMARY

Five well tamed, mature rhesus monkeys, all having had extensive discrimination testing experience, were trained to make visually guided discriminations in the Wisconsin General Test Apparatus. The learning performance of two normal control animals was compared with that of three brain operated animals. Of the three operated animals, one had undergone unilateral extirpation of the major portion of areas 18 and 19 (Brodman), in another areas 18 and 19 were extirpated bilaterally and the third was subjected to bilateral extirpation of the lateral surfaces of the temporal lobes in addition to areas 18 and 19 bilaterally. The operated animals showed retardation in the learning of stimulus objects differing as to form, dissimilar stimulus patterns, and mirror-image patterns. The operated animals' performance could not be differentiated from that of the normals in the discrimination of mirror-image objects. In general, the extent of loss was related to the extent of the lesion.

The two control animals were subsequently operated and postoperative retention tests on these animals failed to reveal any loss. Postoperative retention tests of the discrimination of colors and of a series of different stimulus objects was nearly perfect.

The fact that postoperative learning showed loss, whereas postoperative retention did not, suggests the interpretation that the extent of preoperative testing experience is a significant variable in evaluating deficits resulting from cortical assault. Such an hypothesis permits an integration of apparently conflicting prior results.

REFERENCES

1. ADES, H. W. Effect of extirpation of parastriate cortex on learned visual discriminations in monkeys. *J. Neuropath. & Exp. Neurol.*, 1946, **5**, 60-65.
2. ADES, H. W., & RAAB, D. H. Effect of preoccipital and temporal decortication on learned visual discrimination in monkeys. *J. Neurophysiol.*, 1949, **12**, 101-108.
3. BONIN, G. VON, GAROL, H. W., & McCULLOCH, W. S. The functional organization of the occipital lobe. *Biol. Symp.*, 1942, **7**, 165-192.
4. HARLOW, H. F. Response by rhesus monkeys to stimuli having multiple-sign values. In McNEMAR, Q., & MERRILL, M. A., (Eds.), *Studies in Personality*. New York: McGraw-Hill, 1942.
5. ———. Studies in discrimination learning by monkeys: I. The learning of discrimination series and the reversal of discrimination series. *J. Gen. Psychol.*, 1944, **30**, 3-12.
6. ———. Studies in discrimination learning by monkeys. IV. Relative difficulty of discrimination between stimulus-objects and between comparable patterns with homogeneous and with heterogeneous backgrounds. *J. Gen. Psychol.*, 1945, **32**, 317-321.
7. ———. The formation of learning sets. *Psychol. Rev.*, 1949, **56**, 51-65.
8. KLÜVER, H., & BUCY, P. C. An analysis of certain effects of bilateral temporal lobectomy in the rhesus monkey with special reference to "psychic blindness." *J. of Psychol.*, 1948, **4**, 33-54.
9. ———. Preliminary analysis of functions of the temporal lobes in monkeys. *Arch. Neurol. & Psychiat.*, 1939, **42**, 979-1000.
10. LASHLEY, K. S. The mechanism of vision: XVIII. Effects of destroying the visual "associative areas" of the monkey. *Genet. Psychol. Monog.*, 1948, **37**, 107-166.

Department of Psychology
Emory University
Emory University, Georgia

Department of Anatomy
Emory University
Emory University, Georgia

THE ASSUMPTION OF MOTOR PRIMACY AND ITS SIGNIFICANCE FOR BEHAVIORAL DEVELOPMENT*

Department of Psychology, University of Connecticut

W. A. BOUSFIELD

A. PURPOSE

This paper undertakes to consider what is here termed *the assumption of motor primacy*. It may be stated briefly as follows: *The neuro-motor mechanism develops in advance of capacity to respond to stimulation of sense organs*. The basis of this formulation is in Coghill's (5) monograph on the development of amblystoma. Our interpretation of evidence, however, goes beyond Coghill. We shall attempt to indicate the usefulness of the assumption and its correlates in explaining certain characteristics of behavioral development, and to demonstrate the application of this approach to relevant theories.

B. EVIDENCE OF MOTOR PRIMACY

Coghill (5) undertook to connect his observations on behavioral development with the results of a careful study of the growth of the embryonic nervous system. During the non-motile stage of amblystoma muscular reactions can be elicited only by direct stimulation. At this time both motor and sensory cells develop contact with their respective organs, but the lack of commissural connections prevents responses to sensory influx. The connecting cells are initially unipolar and extend axons to the motor cells on one side only. Later the central neurons become bipolar and establish contact with sensory axons on the opposite side of the cord. Such closing of reflex gaps progresses generally in a cephalocaudal direction. Axons are initially *naked protoplasmic threads*, and their lack of myelinization is associated with considerable diffusion of response. This description provides an explanation for the early flexure and coil stages in which a light touch to the anterior skin elicits first a movement of the head away from the side stimulated and later a cephalocaudal progression of bending down the entire side of the embryo. It would thus appear that during the early development of amblystoma the number of stimuli capable of eliciting motor responses is limited, and that reactions are apparently general in character. We need not

*Received in the Editorial Office on July 18, 1951.

be committed, however, to the generality of Coghill's total-pattern theory. As long as there are grounds for supposing the closure of simple reflex gaps, relatively specific response patterns are possible though they may be components of mass action. The studies undertaken in the past 20 years appear to confirm the developmental priority of the motor system. Hooker (10), in a recent summary of evidence, reports general agreement on several facts of development. Included in these are the conclusions that the capacity to transmit nervous impulses develops in motor nerves prior to the sensory, and that the last components of the reflex arc to become functional are the sensory and very likely the intercalated elements. Hooker further indicates the probability that proprioceptive reflexes appear earlier than the exteroceptive.

We have implied that responses other than those elicited by direct stimulation of muscles depend on activation of sense organs. There is evidence, however, for supposing such an interpretation to be too narrow. Weiss (17) demonstrated the occurrence of uncoordinated motor activity in the transplanted limb of a salamander larva when it received innervation from a deplanted section of spinal cord. There were no afferent fibers and the motor activity must have been due to spontaneous firing of cells in the fragment of spinal cord. Hebb (8) and Weiss (18) cite evidence from a variety of sources indicating the occurrence of spontaneous firing, and such autonomous activity is apparently an established property of aggregates of neural cells. This suggests that a basis is provided for neurally instigated motor responses that may occur without afferent influx.

Evidence significant for the assumption of motor primacy appears in the order of structural development of the human brain. The motor areas appear to develop in advance of other centers. Bolton and Moyes (2) undertook a study of cellular differentiation of the cortex of a fetus 18 weeks of age. They found that the pyramidal cells in the frontoparietal region showed the most advanced development. This part of the cortex is the beginning one of the major motor pathways extending to the spinal cord. McGraw (12) has reviewed this and other studies which similarly indicate relative precocity in the development of motor areas.

If our interpretation of evidence is valid, we may suppose that motor primacy is involved in the following correlated mechanisms: (a) Idio-muscular response consequent on direct stimulation of muscles. (b) Diffuse motor response consequent on sensory stimulation capable of inducing diffusion of excitation in motor tracts. (c) Spontaneous motor response consequent on the firing of cells in the central axis. (d) Isolated reflex response consequent

on localized stimulation and mediated by a neural arc. (e) Concatenated response pattern consequent on the orderly coördination of two or more neuro-motor units.

Except for the priority of idio-muscular response it is not intended that this enumeration should represent a strict temporal sequence. The treatment so far presents a general outline of the evidence indicating the developmental primacy of the neuro-motor mechanism. We shall now consider the usefulness of the foregoing assumptions in accounting for observed facts relating to behavioral development. An attempt will also be made to indicate certain of the variables having a bearing on the interpretation of experimental findings.

C. INTERPRETATION OF ASPECTS OF DEVELOPMENT

1. *Spontaneous Behavior*

Various investigators of behavioral development during the fetal period have noted the occurrence of behavior in the absence of imposed stimulation. Carmichael (4) reported spontaneous movement as the first observable response in his study of the development of fetal guinea pigs. This movement comprised a lateral flexion of the neck accompanied by movement of the forelimbs. In a further study, Bridgman and Carmichael (3) indicated the occurrence of spontaneous behavior 10 to 14 hours after responses elicited by stimulation. Hooker (9) recorded the occurrence of spontaneous movements in the human fetus at the age of $9\frac{1}{2}$ weeks. Responses to tactual stimulation were possible, however, at the earlier age of $8\frac{1}{2}$ weeks. Windle (20) concluded that the time of occurrence of spontaneous activity of the mammalian embryo within the uterus is unknown. However, he inferred that such activity probably occurs appreciably later than movements elicited by stimulation.

Available evidence thus indicates the early occurrence of motor activity in the absence of identifiable stimulation; generally the term spontaneous is used to describe such behavior. That such responses appear to be somewhat spasmodic in character is not surprising since the kinesthetic mechanisms permitting coördinated activity are relatively slow to develop. The study of Weiss (17), to which reference has already been made, reports the occurrence of motor activity as a consequence of the firing of motor cells. We may venture the supposition that this type of neural activity accounts for behavior which should be labelled spontaneous. The time of occurrence of spontaneous behavior would depend on an optimum development of the cells of either motor or correlator neurons. What constitutes such optimum

development cannot at the present time be specified. We do, however, have an important suggestion relating to the rate of growth of segmental neural systems, and Barron (1) has assembled evidence bearing on this problem. He points out that the growth of the limb parallels the growth of cells in the basal plate of associated regions in the central axis. There is an apparent tendency for limbs to appear relatively earlier in some animals than in others. The relative timing of development in axial and appendicular systems has a bearing on the appearance of local reflexes. If our analysis is valid we would suppose that with precocious limb development we should expect a relatively early appearance in the limbs of both reflex and spontaneous responses. The precise estimation of timing of the occurrence of spontaneous response may well depend on use of adequate instruments to detect action potentials in various muscles under conditions which exclude the possibility of sensory stimulation.

2. Motor Diffusion

The term motor diffusion is here employed to denote the essential variability, generality, and relative lack of gradient of response occurring in early development. For example, Minkowski (13) observed that a tactual stimulus applied to the skin of the 135 mm. human fetus induced various reactions in different parts of the body. He observed that every section of the skin could serve as a reflexogenous zone, and that there was a tendency for the reactions to spread over the entire organism. The studies undertaken since Minkowski generally indicate the presence of diffuse and variable responses in the early stages of development. The issue of controversy appears to relate primarily to the relative proportion of the total responses that should be labelled isolated, discrete, or genuinely reflex in character. It is not surprising that motor diffusion appears also as a consequence of direct cortical stimulation. Fulton (6) summarized evidence showing that isolated motor responses are difficult to obtain in the infant macaque. He believes this to be a consequence of the relatively undifferentiated character of the Betz cells and the lack of myelination of the pyramidal tracts. The same phenomena appear also from work with human infants.

The assumption necessary for an explanation of motor diffusion may be indexed in general terms. Motor primacy indicates a relatively rapid development in motor pathways. Myelination of neurons comprising these pathways proceeds as an aspect of their structural development. Whether or not we can regard the myelin sheath as an insulating agent preventing the overflow of impulses, it appears that lack of myelination is associated with

lack of specificity of response. We also assume that afferent pathways from sense organs are relatively slow in establishing functional central connections. If this analysis is valid, it would follow that during the early stages of development, a limited number of stimuli, especially when they are intense, would be effective in eliciting diffuse motor responses. It should be noted further that evidence of motor diffusion would also be expected in spontaneous responses.

3. *Dominance of Intrinsic Neural Activity*

Gibbs and Gibbs (7), in their survey of evidence on developmental changes in *EEG*, indicate the occurrence in newborn infants of large slow waves of one-half to two cycles per second. Waves of this type, according to Hebb (8), are also characteristic of sleep or coma in older subjects, and are probably a result of hypersynchrony in the firing of cortical cells. Hebb suggests that the behavioral consequences of this hypersynchrony include a raised threshold for sensory stimulation, and a lack of the maintained and directed activity which generally is labelled voluntary or purposeful. Stimulation of receptors not only breaks up hypersynchrony, but is also necessary for the co-ordination and control of behavior. At birth intrinsic organization of cortical activity is dominant, and such organization is opposed to the organization consequent on sensory events. The assumption of motor primacy receives support from Hebb's interpretation, but at the same time suggests implications going somewhat beyond his analysis. The early deficiency in central connections with sense organs should contribute both to the heightening of sensory thresholds and to the maintenance of hypersynchrony. It may further be assumed that hypersynchrony is likely to result in the instigation of impulses over motor pathways, and that this type of intrinsic neural activity may sometimes be responsible for spontaneous responses.

4. *Early Learning*

Our discussion of early learning will center on evidence from human fetal and neonatal subjects. It appears that even at birth there is still a lag in the capacity to respond to stimulation, and that observable responses reflect the consequences of diffusion. It would follow that the response specificity required for the conditioning of specific reflexes would be more difficult to obtain at this time than at a later period. On the other hand, since diffuse responses are elicited with comparative ease during the fetal period, the conditioning of such responses should be possible at a relatively early age.

The requirements for such conditioning would involve the use of: (a) an unconditioned stimulus capable of eliciting diffuse responses; (b) a conditioned stimulus which would not elicit diffuse responses. It would be unnecessary that the response to the conditioned stimulus be recordable. As long as conditioning could be established, the requirements for learning would be fulfilled.

Our survey of the literature yields evidence consistent with these deductions. As Wenger's (19) extensive study has shown, the conditioning of specific reflex responses, e.g., lid closure to tactual vibration, in the neonate is difficult. Furthermore, the learning, in so far as it is indicated, is unstable. It appears he was able to establish the conditioning of some responses, but not in all his subjects. Among the difficulties he reported were those involved in the choice of both unconditioned and conditioned stimuli. Many stimuli elicited no observable responses, and the responses he was able to obtain tended to be diffuse and variable. Wenger's findings appear to be in accord with the conclusions reached by Pratt (14) in his survey of studies of neonate learning.

Evidence relating to the conditioning of general activity, while limited, appears to be considerably more positive. In the present frame of reference, the general activity characteristic of the fetus and neonate is regarded as involving diffuse motor responses. Spelt (16) reported success in an experiment designed to condition such responses in the fetus during the last two months of gestation. Employing a procedure suggested by Ray's (15) early work, he used for his unconditioned stimulus a loud noise capable of eliciting fetal movements. This was combined with a vibrotactile conditioned stimulus. Three pairs of tambours attached to the mother's abdomen recorded the presence or absence of fetal movements in the general areas of the head, arms, and legs. The protocols also included signals given by the mother of felt movements. It was not expected that the records would furnish data on the comparative motility of the separate parts of the fetus. In addition to establishing conditioned responses, Spelt also reported success in demonstrating experimental extinction, spontaneous recovery, and retention over an interval of three weeks.

Marquis (11) was able to establish the temporal conditioning of general activity in neonates. A group of infants was fed on a 3-hour schedule for the first eight days, and then shifted to a 4-hour schedule. Stabilimeters were used to obtain records of activity. The results showed adaptation to the 3-hour schedule, and an upset when a change was made to the 4-hour schedule. Furthermore, the activity cycles of the 3-hour group differed

markedly from those of controls fed on a 4-hour schedule, and from those of a group placed on a self-schedule determined by their crying.

If our interpretation is valid, evidence of early learning tends to confirm the assumption of motor primacy and the correlated factor of diffusion. Such learning is limited on the one hand by the lag in responsiveness to stimulation, and on the other by the lack of response specificity.

Mention may appropriately be made here of the relevance of our assumptions to Hebb's (8) concept of the cell assembly and his account of early learning. Primary learning, according to Hebb, involves the establishment of environmental control over groups of cells in the cortex and diencephalon through the building up of cell assemblies. These assemblies are necessary for the establishment of perceptual elements which must precede the emergence of complex perceptions. The formation of an assembly depends on the repetition of particular stimulation and on some degree of constancy of central action following such stimulation. Hebb calls attention to the slowness of this process even for simple perceptions. In relating Hebb's postulates to those of this paper, it may be said that the development of cell assemblies is retarded both by the limited nature of early sensory impact, and by diffusion which works against specificity of response in cortical association areas.

5. *Theories of Development*

Two general theories, essentially opposite in their implications, have emerged from the studies of early behavioral development. This opposition is between the *total-pattern* and the *isolated-reflex* points of view. Coghill (5) was a total-pattern exponent. Development of behavior, he believed, involves the expansion of a general pattern which emerges as a perfectly integrated unit. Partial systems develop within the total pattern and acquire more or less specificity. Against this Gestaltic conception, and after extensive researches, Windle (20) continues to subscribe to the point of view that early movements elicited by stimulation are reflex in nature and depend on functional reflex arcs. His point of view indicates that with progressive development of central reflex mechanisms, more complex and integrated patterns make their appearance. The so-called mass-activity observed by many investigators who have removed fetuses from the uterus is, he assumes, a consequence of asphyxia. During such a state, a diffuse discharge of most of the motor units may be induced by stimulation or may occur spontaneously.

Since, as already indicated, the appearance of specific reflex responses is related to the timing of the emergence of appendicular systems, it might seem

attractive to suppose that the applicability of the total-pattern and isolated-reflex explanations would depend on the type of organism being considered. Several difficulties, however, are apparent in such an approach. Even in mammalian fetuses where appendicular development is relatively precocious, there is no general agreement on the validity of either the isolated-reflex or total-pattern explanations. Further, it is apparent that since such explanations have generated confusion, we should re-examine the problem in order to determine useful and experimentally verifiable assumptions. Nearly 20 years ago Carmichael (4) stated that the formulation of general theories of development such as those denoted by the terms individuation and integration is premature. His observation is still valid.

D. SUMMARY AND CONCLUSIONS

It is obvious that the incidence of various forms of behavior must depend on the physiological mechanisms available to the organism, and the study of behavioral development should be guided by relevant premises derived from developmental physiology. Present evidence supports the assumption of the developmental priority of the neuro-motor system, and a corresponding lag in the capacity to respond to stimulation. The consideration of this assumption has led to the supposition of correlated mechanisms for the eliciting of muscular responses. This approach has been shown to be useful in accounting for various aspects of behavioral development, and at the same time furnishes a basis for the evaluation of relevant theories.

The literature shows considerable difference of opinion with respect to the facts of behavioral development and a consequent lack of agreement in theory. We would propose that the resolution of this problem depends in large measure on adequate definitions of the behavioral phenomena available for observation, and that such definitions should derive from tenable assumptions relating to the development of physiological mechanisms for response. Our analysis has led to the postulation of motor primacy and the definitions of five correlated mechanisms of motor reaction. If these assumptions are valid, there should be no essential disagreement on the facts they denote, and facts are necessary for both the generation and the testing of theory. As a general orientation for further study, we would suggest the following:

1. *Idio-muscular response.* More information is needed on the time when muscles are capable of responding to direct stimulation, the nature of such responses, and the types of effective stimuli.
2. *Diffuse motor response.* Motor diffusion denotes the occurrence of

non-specific responses as a result of the spread of excitation over neural pathways. There is a need for more effective measures of this phenomenon and the conditions on which it depends. The latter would appear to include the physiological variables of the level of maturity of motor pathways, and the condition of the organism. Research is needed on the nature and locus of receptor stimulation capable of eliciting this type of response, and on the occurrence of diffusion as an aspect of spontaneous behavior. It would also be worthwhile to identify sequences in the spread of excitation.

3. *Spontaneous response.* Such responses would appear to depend on the firing of neural cell aggregates in either lower or higher centers. Further research might well be undertaken to determine the occurrence of such responses following the elimination of afferent influx of impulses from receptors.

4. *Isolated reflex response.* By definition the reflex is a constant type of response following the application of a specific stimulus. Such an element of behavior is assumed to depend on a neural arc connecting specific receptors with specific motor or glandular organs. The identification of isolated reflex responses is complicated, however, by the consequences of diffusion occurring in the central axis and presumably also in peripheral nerves. To be labelled as a specific reflex response, however, the behavior should be mediated by a functioning neural arc. Consideration should be given to the consequences of development occurring within its components.

5. *Concatenated response pattern.* As is the case with other types of response, it is apparent that the problem of identification is complicated by the occurrence of diffusion. There is no sharp transition between generalized and indeterminate responses on the one hand, and an orderly coördination of multiple responses on the other. It would seem necessary, however, to specify the components of the pattern being studied, and further to indicate their temporal relationships.

REFERENCES

1. BARRON, D. H. Genetic neurology and the behavior problem. In (P. Weiss, Ed.) *Genetic Neurology*. Chicago: Univ. Chicago Press, 1950. Pp. 239.
2. BOLTON, J. S., & MOYES, J. M. The cyto-architecture of the cerebral cortex of the human fetus of eighteen weeks. *Brain*, 1912, **35**, 1-25.
3. BRIDGMAN, C. S., & CARMICHAEL, L. An experimental study of the onset of behavior in the fetal guinea-pig. *J. Genet. Psychol.*, 1935, **47**, 247-267.
4. CARMICHAEL, L. An experimental study in the guinea-pig of the origin and development of reflexes and patterns of behavior in relation to the stimulation of specific receptor areas during the period of active fetal life. *Genet. Psychol. Monog.*, 1934, **16**, 338-491.

5. COGHILL, G. E. *Anatomy and the Problem of Behavior*. Cambridge: Univ. Press, 1929. Pp. 113.
6. FULTON, J. F. *Physiology of the Nervous System*. New York: Oxford Univ. Press, 1943. Pp. 614.
7. GIBBS, F. A., & GIBBS, E. L. *Atlas of Electroencephalography*. Cambridge, Mass.: Cummings, 1941. Pp. 221.
8. HESS, D. O. *The Organization of Behavior*. New York: Wiley, 1949. Pp. 335.
9. HOOKER, D. Reflex activities in the human fetus. In (R. G. Barker, J. S. Kounin & H. F. Wright, Eds.) *Child Behavior and Development*. New York: McGraw-Hill, 1943. Pp. 652.
10. ———. Neural growth and development of behavior. In (P. Weiss, Ed.) *Genetic Neurology*. Chicago: Univ. Chicago Press, 1950. Pp. 239.
11. MARQUIS, D. P. Learning in the neonate: The modification of behavior under three feeding schedules. *J. Exp. Psychol.*, 1941, **29**, 263-282.
12. MCGRAW, M. B. Maturation of behavior. In (L. Carmichael, Ed.) *Manual of Child Psychology*. New York: Wiley, 1946. Pp. 1068.
13. MINKOWSKI, M. Über frühzeitige Bewegungen, Reflexe, und muskuläre Reaktionen beim menschlichen Foetus, und ihre Beziehungen zum fötalen Nerven- und Muskelsystem. *Schweiz. med. Wschr.*, 1922, **52**, 721-724, 751-755.
14. PRATT, K. C. The neonate. In (L. Carmichael, Ed.) *Manual of Child Psychology*. New York: Wiley, 1946. Pp. 1068.
15. RAY, W. S. A preliminary report on a study of fetal conditioning. *Child Devel.*, 1932, **3**, 175-177.
16. SPELTZ, D. K. The conditioning of the human fetus in utero. *J. Exp. Psychol.*, 1943, **38**, 333-346.
17. WEISS, P. Autonomous versus reflexogenous activity of the central nervous system. *Proc. Amer. Phil. Soc.*, 1941, **84**, 53-64.
18. ———. An introduction to genetic neurology. In (P. Weiss, Ed.) *Genetic Neurology*. Chicago: Univ. Chicago Press, 1950. Pp. 239.
19. WENGER, M. A. Conditioned responses in human infants. In (R. G. Barker, J. S. Kounin & H. F. Wright, Eds.) *Child Behavior and Development*. New York: McGraw-Hill, 1943. Pp. 652.
20. WINDLE, W. F. Reflexes of mammalian embryos and fetuses. In (P. Weiss, Ed.) *Genetic Neurology*. Chicago: Univ. Chicago Press, 1950. Pp. 239.

Department of Psychology
University of Connecticut
Storrs, Connecticut

FOOD CONSUMPTION AND GROWTH OF THE RAT AS A MEASURE OF MOTIVATIONAL STRESS*¹

Western Psychiatric Institute and Clinic, University of Pittsburgh

FRANCIS J. PILGRIM AND R. A. PATTON²

A. INTRODUCTION

Motivational stress or "emotional conflict" has led to behavior abnormalities in several species of animals. A number of methods in the field of comparative psychopathology have been found effective in producing the so-called "experimental neuroses." Some of the procedures found most effective have imposed upon the animal an increasingly difficult discrimination in conditioning or learning situations. Others have depended on frustration or conflict of basic biological drives. Studies by Pavlov (22, 23), Liddell (16), Gantt (9, 10), Masserman (19), and others have emphasized methods for producing and measuring such signs as well as descriptions of the progressive involvement of different physiological systems. Although Pavlov introduced the term "experimental neurosis," it would be better objectively to describe the signs appearing in animals as behavior abnormalities (28).

Shenger-Krestovnikova, in Pavlov's laboratory (22, Ch. 17), produced an experimental neurosis in dogs by using difficult discriminations of stimuli for conditioned responses. Another experiment (22, Ch. 18) illustrated the effect of a rapid increase in strength of the conditioned stimulus in a delayed reaction problem. The positive conditioned reflexes disappeared and later some of the inhibitory stimuli began to give positive effects. Studies of disturbances of behavior in dogs have been extended by Gantt (9). He presented case histories of three dogs and then analyzed the objective signs according to physiological systems. The details are presented under the following major divisions: (a) muscular activity and metabolism, (b) gastrointestinal symptoms, (c) respiratory symptoms, (d) cardiac symptoms, (e)

*Received in the Editorial Office on July 24, 1951.

¹This paper reports research undertaken in cooperation with the Quartermaster Food and Container Institute for the Armed Forces, and has been assigned number 319 in the series of papers approved for publication. The views or conclusions contained in this report are those of the authors. They are not to be construed as necessarily reflecting the views or indorsement of the Department of Defense.

²The authors are indebted to Francis H. Palmer for assistance in conducting these experiments.

urinary symptoms, (*f*) sexual symptoms, (*g*) social behavior, (*h*) generalization of the neurotic behavior, (*i*) clinical correlations. It should be noted particularly that when the abnormal behavior developed, the animals refused food in the experimental situation, and also in other environments they refused the food that had been used in the experimental studies. Maier (17) also found that rats in an unsolvable discrimination problem refused food in the apparatus, but not in their living cages.

Liddell (16) and his co-workers have extended the conditioned response technique to sheep and pigs as well as performing studies on dogs. They have also studied the respiratory and cardiovascular responses. Anderson and Parmenter (1) have pointed out two main factors in animal neuroses. The one, a difficult discrimination, has been emphasized by a number of investigators; further, they ascribed part of the effect to a factor ordinarily present in conditioning problems, a monotonous routine (1, pp. 71-72). Rhythmical stimulation alone in the dog, goat, and sheep produce and maintain a neurosis. No problem is involved, other than repeated positive conditioned stimuli. They speak of the "deadliness" of the temporal routine. Animals are able with fair accuracy to time an interval between the stimuli, hence "the waiting time interval itself becomes a conditioned stimulus efficient enough to evoke . . . conditioned defensive movements of the reaction limb and head, distorted and hastened respiration and rapid pulse—the chief signs of the experimental neurosis in the laboratory."

Masserman (19, p. 117) has summarized the conditions that are effective in producing experimental neuroses by the conditioned reflex method. These are: (*a*) when a weak conditioned stimulus is rapidly made strong; (*b*) when strong or protracted inhibition is imposed; (*c*) by difficult or monotonous differentiation; (*d*) by employing painful stimuli; and (*e*) by restraint of voluntary movements. Cook (4) has summarized the consequences of some of these earlier experiments.

The report of Masserman (19) on cats, although employing conditioned responses, was based mainly on frustration or conflict of drives, as was that of Maier (17, 18). These studies reported the behavioral changes observed. Gantt (9) has pointed out that other factors may be involved since animals in natural environments do not usually develop abnormalities from conflict of drives such as food-getting and pain-avoidance. He suggested that in fighting for food the aggression is active with free and uninhibited expression, whereas the laboratory situation involves passive, restrained, and unsuccessful defense. "Frustration is present consistently and repeatedly in one situ-

ation, but not in the other" (19, p. 34). Masserman presented methods of inducing frustration and conflict, and further stated that the most significant experiments were those in which a conflict of incentives was set up, e.g., patterns of feeding at a box were made incompatible with those of escape from a blast of air across the box just as the food was delivered (19, p. 27). He thus obtained motivationally conflictful or "neurotic" behavior.

There are few reports of abnormal behavior in rats. Of these, the interpretation of one has been contested and some of the others employing few rats has not been confirmed. The abnormal behavior attributed by Maier (17) to conflict has been shown by other investigators to be essentially a sound-induced convulsion, and hence has little bearing on the present problem [see the review by Finger (8)]. However, Maier's method of inducing conflict on the Lashley discrimination apparatus (15) may prove to have value in the study of behavior. Studies of abnormal fixations in the rat have been reported (18), using this technique. A report (21) of enduring emotional responses in the rat from intermittent electric shock could not be confirmed (6). This point will be considered further in the experimental section of the present paper. Cook (5), and also Humphrey and Marcuse (12) have described a few instances of so-called experimental neuroses in rats. The former observed such behavior in three out of six rats when conditioning a leg flexion in a restricting stand. He stated that activity must be limited to the response situation. The other study (12) produced behavioral changes by moving the food box at the end of a maze after the rat had entered it. The box was on the floor and thus forced the rat to walk inside of it, as it was moved across the floor. This resulted in tenseness, gnashing of teeth, and refusal to enter the box at the end of the maze.

There is little in the above-mentioned reports that would indicate that any severe or long-lasting "experimental neurosis" had been induced in rats by laboratory procedures. Hunt and Schlosberg (13) reported a study in which the drinking tubes in the rats' cages were electrified. They observed decrements of activity, and when the electricity was turned off, behavior reminiscent of that associated with hallucinations in humans, and short bursts of activity. This lasted only a few days after the shock was discontinued. They suggest that the rat perhaps does not have the symbolic capacity "to allow it to carry its troubles about" (7).

It seems necessary to set objective criteria of measurement of behavioral changes with systematic long-term experiments to determine to what extent rats exhibit abnormal behavior. Hebb has attempted to clarify some of the problems involved (11).

Many of the studies cited above have used food for motivation with punishment (pain, "fear") or withholding of the food as the conflict or frustration. Marked changes in the behavior of the animals toward food were often noted. Cannon (2, 3, pp. 434-448) has demonstrated the close relationships existing between emotional stress in animals and impairment of the digestive processes. There is the possibility that lowered efficiency in meeting continued stressful situations may have been due to a prolonged interference with normal digestion and nutrition. This is particularly true when partial starvation is employed to intensify the drive. It has more recently been shown that starvation increases the intensity of the abnormal behavior in animals (29).

No clear-cut attempt has yet been made to relate such emotional disorders to nutritional factors. Do experimental animals subjected repeatedly to motivational stress show alterations in their behavior towards food? Are such changes confined to the experimental situation or do they continue when the animals are replaced in their living quarters?

Stress, in the present problem, is a condition or situation that requires an organism to attempt to adapt or readjust its physiology or behavior to compensate for internal or external changes which cannot be avoided. These factors may influence the responses at two levels of activity: one at the automatic or physiological level such as the increase in metabolic rate to compensate for changes in temperature; the other at the level of learned activity involving more or less specific behavior patterns such as food-seeking or avoidance of painful situations. Automatic or physiological changes may be present along with the learned or behavioral activity.

The stress employed in the experiments discussed above had, in general, been for short periods each day. This apparently was effective in the species used. However, in an animal such as the rat it may be necessary to subject them to long-enduring stressful situations in order to obtain behavioral abnormalities. Previously, no such experimental designs have been reported that have systematically explored the effect of protracted conflict on rats.

B. THE PROBLEM

The main aim of the investigations to be described has been to study systematically whether rats can be subjected to one or more of several types of motivational stress that will alter their growth rate, total food consumption, or patterns of selection of the purified components of the diet. These criteria were chosen because of their objectivity. The rat is the animal of

choice because of its uniformity of response in physiological measurements, and because of the common genetic background of the highly inbred strains that are available. Diets composed of purified components have been employed in order to control nutritional factors and to have a more quantitative measure of changes in food intake. The purified diet has the further advantage that the consumption of the individual components can be measured in a free choice situation while the rats are subjected to stress.

In the studies there are three factors involved: (a) duration or time: whether the stress or stimuli last for short or long periods each day; (b) activity-passivity of the animal; whether the rat initiates the stress or stimulus by performing an act to satisfy a drive or submits passively to the stimulation as regulated by the experimenter; (c) hunger (or thirst): whether food is the basis for one of the drives or provides the motivation for a particular act. The latter could be used only in an "active" situation; i.e., when there is conflict between hunger and some other drive. A "short-passive" type of experiment; i.e., stress from inability to escape from noxious stimuli, would be desirable only if the "long-passive" type produced measurable changes. The "short-active-hunger" type requires long periods of starvation, and hence would be undesirable from the nutritional standpoint. A preliminary experiment of this type was made, but the subsequent one was changed to non-hunger. Of the possible types of situations, three are reported herein.

1. *Short-Active-Non-Hunger*

The rat is forced to perform in an undesirable situation for a short period each day. In this experiment the animal is taught a discrimination problem on the Lashley apparatus (15) and then the problem is made unsolvable as described by Maier (17). Avoidance of shock is used as the motivation so that food is not involved.

2. *Long-Passive-Non-Hunger*

Noxious stimuli of several types occur at irregular intervals as predetermined by the experimenter, and are independent of anything the rat does. Shock, air blast, and a loud electric buzzer have been chosen as the stimuli. These occur every few minutes for 24 hours a day. The rat cannot escape the situation.

3. *Long-Active-Hunger*

In this experiment, an attempt to develop a conflict is made by having the rat turn on an electrified grid in front of the food container each time

it attempts to eat. Opposing drives, that to obtain food, and that to avoid pain, are present simultaneously.

4. *Long-Active-Non-Hunger*

This has been investigated in a preliminary experiment only and will not be presented in this report. The situation is based on a conditioned response to pain avoidance in which there is no escape. A temporal delay is introduced and gradually lengthened in order to produce an expectancy of the pain that cannot be avoided. Mowrer (20) has devised a situation that might also prove useful for the study of this type of frustration.

C. EXPERIMENTAL

1. *Animals and Diets*

In all experiments in the present study the same purified diets and the same type of rats were employed. The details relating to these will be presented once before the discussion of the individual experiments.

In all the experiments young, growing rats were used. The rat was chosen because of its uniform physiological responses and common genetic background. Male, albino rats of the Sprague-Dawley strain, all of the same age, 21 days, and within a weight range of 10 grams, weighing 35-45 grams at the start, were employed for each of the experiments. Food consumption and growth were measured regularly throughout the course of each study.

In order to control nutritional factors and to be able to carry out studies on self-selection of diet, a purified ration of known composition was employed. Casein was used as the sole source of an adequate protein. The diet had the following composition, in percentages: casein (Labco, "vitamin-free") 18; sucrose, 73; salts 12 (14), 4; fat (hydrogenated vegetable oil), 5. Each 10 grams (the approximate minimum daily consumption of the rat) contained the following quantities of vitamins: thiamine hydrochloride, 40 μ g.; pyridoxine hydrochloride, 40 μ g.; calcium pantothenate, 200 μ g.; niacin, 200 μ g.; choline chloride, 10 mg.; vitamin A, 100 I.U.; vitamin D, 20 I.U.; α -tocopherol, 200 μ g.; and 2-methyl-1, 4-naphthoquinone, 40 μ g.

In the studies of self-selection, the four major constituents were presented to the rats separately and the vitamins were given as a supplement in a tablet of sucrose-dextrin base.

2. *Experiment 1 (Type: Short-Active-Non-Hunger)*

The earlier studies (17, 18) that used the Lashley discrimination apparatus to induce conflict employed food reward as the motivation. In

order to use shock avoidance as the motivation it was felt desirable to make several preliminary studies, first using food as motivation, and then using shock, so that the two techniques could be compared.

After the study employing the hunger drive, the apparatus was modified so that shock of a graded intensity could be applied to the rat's paws to force him to jump. The results in terms of learning the discrimination were similar. It was also found that the resistance of the animal was great enough in the food reward situation when conflict was introduced that it was necessary to use prodding or a form of shock (inductorium) to force him to move. Because of these factors the details of only the major experiment on the rôle of this form of conflict in growth and food consumption will be presented.

a. Animals. Twenty-four rats were housed in individual wire mesh cages and started on the diet at 22 days of age. The rats were weighed four times per week.

b. Food and water consumption. The rats were fed the purified diet described above, and water *ad libitum*. The food was placed in one-ounce cups with a shield-type lid to minimize spillage. Spilled food was caught on paper below the screen bottom of the cage so it could be recovered when the food consumption was measured. Water was supplied in graduated drinking tubes. Both the food and water consumptions of the individual rats were measured daily.

c. Apparatus. The Lashley discrimination apparatus (15) consists of two adjacent windows each 6 x 6 inches in a vertical black panel with a table behind the windows. Cards of a given form could be inserted in the windows. In front of the windows is a jumping platform. The platform in the present device was modified from the original and hence will be described in detail. This consisted of a box 4 inches wide by 6 inches long by 6 inches high. The floor was a grid of monel metal bars $\frac{1}{8}$ inch in diameter, spaced $\frac{1}{2}$ inch apart. This could be electrified with 0 to 110 volts by an adjustable potentiometer. The animals were induced to jump towards the cards in the windows by being given a shock of one or two seconds duration every five or ten seconds. The voltage was started near the sensory threshold and increased gradually, five or ten volts at a time, until the rat jumped off the platform. No food was placed on the table behind the windows, because the animals were not restricted in their food intake at any time.

d. Procedure. In a preliminary experiment, 18 rats were taught the problem and then subjected to the conflict. The results indicated that the

modified procedure was satisfactory. It was then used with a series of young rats whose growth rates and food intakes were measured during the learning and conflict periods.

From the 24th to 28th day the rats were familiarized with the apparatus for five minutes each day. From the 29th to 37th day the rats were run individually with both windows open. The jumping platform was gradually moved back from the windows, until a distance of eight inches was attained. The discrimination cards were introduced on the 38th day. The cards consisted of one black square with a white circle in the middle and one white square with a black circle in the middle. The areas of the white and black portions of each card were equal (18 sq. in.), so that the total reflected light from each card was the same. The rats were trained to knock over the positive or unlocked card to reach the platform behind the cards, and thus escape the shock. For one-half of the rats, the white circle on the black background was made the positive card; for the other half the black circle on the white background was positive. From the 43d day to the end of the experiment each rat was given 10 trials per day at a distance of eight inches from the cards. The pattern for the position of the positive card was altered randomly each day. The rats had learned the discrimination adequately by the 50th day (one rat which refused to jump was discarded). They were then divided into two groups equalized on the basis of weight gain and food consumption, and with one-half of each group having the white circle as the positive card, and the other having the black circle as the positive card. There were 11 animals in the control group and 12 in the experimental or "stress" group. Conflict was introduced on the 51st day by locking the positive window for five of the 10 daily trials. The trials on which the door was locked were partially randomized. The conflict group was continued in this situation to the end of the experiment on the 80th day of age of the rats. After the 60th day the control group was not even required to continue in the discrimination problem. They were just maintained in their cages, and their food consumption and weight gains measured along with those of the experimental group. This should accentuate any differences that might appear.

e. Results. The only overt abnormal behavior of the stress group was a resistance to jumping. It required at times fairly high levels of shock to force them to react to the situation. No unusual behavior was noted in these rats outside of the conflict situation. The data for weight and for the average gain, food consumption, and water consumption for five-day periods (in-

cluding two periods before stress was introduced) are shown in Table 1. For this number of samples the t value necessary for a 5 per cent level of confidence is 2.23; for a 1 per cent level, it is 3.17. It is evident from Table 1 that only in one of the factors measured, that of gain, is there at any time a significant difference between the two groups. This difference is reversed, though not as significantly, during two other periods.

It is concluded from an analysis of the data that motivational stress or conflict of the type used in this experiment did not alter the weight gain or the voluntary consumption of food and water by the rat.

TABLE 1
EFFECT OF STRESS ON WEIGHT, GAIN, AND FOOD AND WATER CONSUMPTION†

Period days*	Group	Weight g.		Gain g./5 days		Food g./5 days		Water cc./5 days	
		M	SE	M	SE	M	SE	M	SE
41-45	Control	149	3.2	23.4	1.06	79	2.4	75	3.2
	Stress	145	2.1	22.3	0.95	81	3.5	72	3.8
	t	1.05		0.77		0.47		0.60	
46-50	Control	174	4.2	24.7	1.14	81	2.7	81	2.7
	Stress	169	3.0	24.7	1.07	83	1.7	71	5.0
	t	0.97		0.00		0.63		1.76	
**51-55	Control	194	4.3	20.0	1.41	83	2.4	87	4.3
	Stress	191	2.5	21.5	0.79	81	3.4	77	4.5
	t	0.60		0.96		0.48		1.60	
56-60	Control	216	4.9	22.4	0.84	83	2.0	94	4.9
	Stress	212	2.5	21.0	0.89	81	2.3	90	5.2
	t	0.73		1.14		0.66		0.56	
***61-65	Control	236	4.8	20.1	0.49	88	1.8	96	4.6
	Stress	230	3.3	17.8	0.67	87	1.9	90	5.6
	t	1.03		2.77		0.38		0.83	
66-70	Control	252	4.7	15.8	0.75	87	2.3	96	5.3
	Stress	244	2.9	13.8	0.68	85	2.3	83	4.6
	t	1.45		1.97		0.62		1.85	
71-75	Control	273	5.4	20.4	1.01	86	2.5	103	4.1
	Stress	260	3.1	16.8	0.85	83	2.2	95	5.8
	t	2.09		2.73		0.90		1.13	
76-80	Control	281	5.4	8.5	1.09	81	2.5	112	6.1
	Stress	269	4.1	8.8	0.92	81	2.0	93	7.9
	t	1.77		0.21		0.00		1.90	

†All data are given in terms of the mean and standard error of the mean. The t value or critical ratio is given beneath each pair of means.

*Days of age of rats.

***"Stress" group started on conflict on Day 51.

***Controls discontinued from discrimination apparatus after Day 60.

3. Experiment 2 (Type: Long-Passive-Non-Hunger)

In the first two experiments of this type not only total food consumption, but also the patterns of self-selection of rats under stress were studied. This necessitated preliminary studies of the patterns of selection in order to obtain homogeneous groups.

In a recent study of the rôle of nutrition in the selection of purified components by the normal, immature rat, a pronounced variability and an abnormal frequency distribution of choices were found (26). Thus, 39 per cent of the animals allowed their choice of casein, salts (complete mixture), sucrose, and fat failed to grow because of an inadequate intake of casein. Aside from the appetite for salts and calories, the appetites of these animals showed no clear-cut relationship to nutritional needs.

The results of our own studies on self-selection have been published (25). They agree in general with the earlier study (26). The summary of the first experiment is presented here before the details of the study of the rôle of stress in patterns of selection are discussed. On self-selection, rats six to nine weeks old demonstrate the following: (a) only 25 per cent gain as well as control rats on a complete purified diet; (b) most rats prefer fat to sucrose; (c) no rat failed to grow when it consumed sufficient casein (a sufficient amount is 1.-1.5 g./day; (d) no rat suffered from a deficiency of salts or of total calories when the protein intake was adequate.

A consideration of these studies indicated that it would be necessary to study the patterns of self-selection in large groups of relatively uniform rats. From these it would be possible to select a homogeneous group of efficient or inefficient selectors. Only then could the effect of introducing an additional variable, such as stress, upon these patterns of choice be adequately appraised.

a. *Animals.* Forty-eight male, weanling, albino rats weighing 35 to 45 grams were employed. They were housed in individual wire mesh cages with screen bottoms and were fed the mixed diets or individual constituents *ad libitum*. Water was available at all times from a drinking tube attached to an inverted bottle. They were fed purified rations or the individual constituents throughout the experiment. At the 35th day of age, when the first period of self-selection was begun, all rats weighed between 73 and 101 grams.

b. *Studies of self-selection.* The rats were allowed to select their own diet from: casein (Labco, "Vitamin-Free"), salts 12, sucrose, and fat (Spry or Crisco). While still on the purified complete diet, the rats were

trained to consume a tablet containing the necessary vitamins.³ Beginning on the 35th day of age, the four constituents were placed in individual cups of one ounce capacity, with a shield-type top to minimize spillage. The cups were equidistantly spaced in holes drilled near the periphery of a board seven inches in diameter. The consumption of each constituent by each rat was measured four times per week (with correction for any spillage), and the cups interchanged in a predetermined random manner to prevent formation of patterns of selection on the basis of position habits.

c. Apparatus. The device was designed to subject the rats passively to stimulation; that is, the long-passive-non-hunger type. It consisted essentially of a large box which was divided into two like parts to provide two group-living quarters. The physical dimensions of each compartment were as follows: 23 inches long by $13\frac{1}{2}$ inches wide by 16 inches high. Three sides were of wood painted black; the fourth side had a glass window for observation. The top was closed by a wire mesh screen to provide ventilation, while at the same time confining the animals. The floor consisted of $\frac{1}{8}$ inch monel metal bars spaced $\frac{3}{8}$ inch center to center. The alternate bars of this grid were attached to the two poles of a voltage divider or potentiometer so that the grid or floor could be electrified with any desired voltage. Along one wall there was a pipe, $2\frac{1}{2}$ inches above the grid, containing three small holes arranged to deliver a blast of compressed air over practically the whole floor area of the compartment. This pipe was connected through an electrically controlled valve (solenoid gas valve) to a source of compressed air at 35 lb./sq. inch pressure. On one wall was mounted a loud electric buzzer. Each of these three stimuli was operated by a separate automatic electric timer set to deliver a 10-second impulse at predetermined intervals.

The four constituents of the diet were accessible to each group of animals at all times, except during the study on starvation. Along the wall of each compartment opposite the air blast there were four apertures, each $1\frac{1}{2}$ inches in diameter with the center $3\frac{1}{2}$ inches above the floor. Containers for the food, with shields above the container and hole, were placed on the outside of the holes. They were so designed that the animals could put their heads through to obtain food without being able to enter far enough to escape from the compartment with its attendant stimuli. A glass tube or drinking fount was inserted through one wall to supply water to the rats.

d. Procedure. On the basis of the patterns of selection for an 18-day period, the rats could be divided into those that selected adequately and those

³This consisted of one 100 mg. tablet of sucrose-dextrin each day containing the quantities of vitamins listed above plus one microgram each of biotin and folic acid (Lederle).

that did not. The inadequate selectors would probably eventually die and hence could not be continued on self-selection in the experiment on stress. The adequate selectors were then subdivided on the basis of whether the calories consumed above the protein necessary for growth (1.0-1.5 g./day) consisted mainly of protein, carbohydrate, or fat. Sufficient numbers of animals that consumed predominantly carbohydrate, or fat were obtained to form two groups of 10 and 8 animals, respectively. One-half of each group was placed in a compartment of the stimulus box and the other halves, the controls, placed in two cages in a different room, away from the noises of the stimulus box. The group patterns of selection of each of these four groups as well as the individual weight records were then followed daily throughout the remainder of the experiment.

The rats in the stimulus box were then subjected to the three stimuli according to the schedule given in Table 2 which lists the chronological sequence of procedures for the whole experiment. In brief, each stimulus occurred about every one to eight minutes for half the day during the first two weeks. For the remaining five weeks of the experiment, the intervals between stimuli were reduced somewhat, and at the same time were continued for the whole 24 hours of each day (Table 2).

The animals were observed frequently for any visible behavioral disorders.

TABLE 2
SEQUENCE OF PROCEDURES FOR EXPERIMENT 2

Day of age	Treatment
24	Rats received and given purified diet in individual cages.
35	Animals placed on self-selection in individual cages.
51	Sucrose group selected and placed in two groups of five rats each.
54	Fat groups selected and placed in two groups of four rats each.
57	Placed groups in stimulus box; controls continued as two groups and treated the same as experimentals except not subjected to the stimuli.
	Stimuli 12 hours a day:
	Buzzer.....3 minute intervals.
	Shock.....5 minute intervals.
	Air-blast.....8 minute intervals.
64	Stimuli 14 hours a day:
	Buzzer.....2 minute intervals.
	Shock.....4 minute intervals.
	Air-blast.....7 minute intervals.
70	Buzzer.....2 minute intervals.
	Shock.....1 minute intervals.
	Air-blast.....5 minute intervals.
78	Stimuli continuous (24 hours a day).
	Intervals between stimuli same as above (Day 70).
89	Food removed for 48 hours.
107	Experiment terminated.

It was also necessary to adjust the voltage supplied to the grid to such a level that the animals gave a noticeable response each time the electricity came on without, however, reacting violently enough to indicate that they were being physically harmed by the current. This voltage varied from 45 to 55 volts from day to day.

e. Results. The values for the average group weight changes and for

TABLE 3
FOOD CONSUMPTION AND WEIGHT GAIN OF STIMULATED AND CONTROL RATS

	Period days of age	Sucrose- stimulated	Sucrose- control	Fat- stimulated	Fat- control*
Weight ²		91.8	86.2	81.2	86.8
Gain		4.5	4.0	3.6	4.0
Casein					
Salts	25-35		On complete purified diet		
Sucrose					
Fat					
Calories					
Weight ²		141.2	139.0	111.8	111.5
Gain		4.0	4.2	5.0	4.4
Casein		3.31	2.61	2.25	2.26
Salts	46-50	0.39	0.38	0.26	0.28
Sucrose		8.33	8.95	0.22	0.49
Fat		0.18	0.32	3.69	3.34
Calories		48.2	47.5	43.1	41.1
Weight ²		177.0	179.2	151.0	155.0
Gain		4.3	4.8	4.7	5.2
Casein		4.54	3.66	3.13	3.47
Salts	56-60	0.86	1.14	0.77	0.62
Sucrose		9.80	9.56	0.24	2.77
Fat		0.27	0.94	3.90	2.80
Calories		59.8	61.3	48.6	50.2
Weight ²		207.0	214.2	192.8	196.8
Gain		1.8	2.6	4.0	2.7
Casein		4.18	3.75	3.20	3.06
Salts	66-70	0.44	0.81	0.49	0.55
Sucrose		9.37	10.31	0.93	10.48
Fat		0.33	0.58	3.75	0.70
Calories		57.2	61.5	50.3	60.5
Weight ²		228.6	241.6	223.0	227.5
Gain		0.8	0.4	1.3	1.5
Casein		3.37	2.68	2.92	2.82
Salts	76-80	0.33	0.65	0.31	0.61
Sucrose		7.87	11.37	0.94	10.96
Fat		0.62	0.28	3.44	0.14
Calories		50.5	58.7	46.4	56.4

*This group changed to predominantly sucrose by Day 63.

TABLE 3 (Cont'd)

	Period days of age	Sucrose- stimulated	Sucrose- control	Fat- stimulated	Fat- control
Weight ^a		238.6	261.6	249.2	243.2
Gain		0.4	1.9	3.0	2.9
Casein		3.42	3.61	3.74	3.62
Salts	86-89	0.19	0.49	0.25	0.28
Sucrose		6.89	9.72	3.64	10.39
Fat		0.25	0.32	1.11	0.25
Calories		43.5	56.2	39.5	58.3
Weight ^a		218.5	243.2	230.8	234.7
Gain	90-91	-9.3	-10.3	-10.2	-9.0
Starved for 48 hours.					
Weight ^a		238.5	260.2	250.8	253.0
Gain		3.6	4.5	4.8	6.0
Casein		4.63	3.82	3.84	3.89
Salts	92-95	0.38	0.54	0.46	0.34
Sucrose		7.83	10.52	3.10	9.72
Fat		0.47	0.39	4.24	0.79
Calories		54.1	60.9	65.9	61.6
Weight ^a		249.7	276.6	269.7	273.5
Gain		2.6	1.6	2.6	1.1
Casein		3.14	2.82	3.58	2.67
Salts	101-104	0.30	0.51	0.26	0.26
Sucrose		7.15	8.97	2.18	9.45
Fat		1.29	0.92	3.14	0.95
Calories		52.8	55.4	51.3	57.0
Weight ^{**}		253.2	277.2	272.8	274.7
Gain		0.6	0.4	-0.1	0.9
Casein		3.59	3.29	3.84	3.09
Salts	105-107	0.28	0.37	0.20	0.30
Sucrose		9.63	11.00	2.68	10.86
Fat		0.58	0.61	3.14	0.33
Calories		58.1	62.7	54.3	58.8

^{**}The weight is for approximately the middle day of the period.

the consumption of the individual constituents of the diet are presented in Table 3. The values for three- to five-day periods have been averaged for convenience in presentation. Some of the periods have been omitted to conserve space. The results were similar to those preceding and following the deleted data. From the values for consumption of casein, sucrose, and fat have been calculated the total caloric consumption using the caloric equivalents of 4, 4, and 9 calories per gram respectively.

The patterns of selection of the rats in the stimulus box did not change; i.e., those that ate predominantly fat during the preliminary measurements

(up to the 57th day) continued to eat fat as the main source of calories; similarly for those originally consuming mainly sucrose. The fat-control group shifted gradually to sugar. This tendency was evident in several of the rats during the preliminary period, but it was necessary to employ them in the experiment anyway because there were too few animals with a completely stable pattern of selection.

The figures for growth rates given in Table 3 would indicate a lower growth rate for the rats in the sucrose-stimulated group. The individual values for the rats are presented in Table 4. From these figures it can be

TABLE 4
WEIGHT AND WEIGHT GAINS OF INDIVIDUAL ANIMALS

Group		Sucrose-stimulated					Sucrose-control				
Rat No.		2	15	19	21	25	5	8	30	33	37
WEIGHT, g	55th day	154	170	167	170	171	155	165	166	163	182
	85th day	218	223	249	254	242	232	255	262	249	291
	105th day	226	245	264	276	*	253	273	288	270	312
GAIN, g	55-85 days	64	53	82	84	71	77	90	96	86	109
	85-105 days**	8	22	15	22	*	21	18	26	21	21
	55-105 days	72	75	97	106	*	98	108	122	107	130
Group		Fat-stimulated				Fat-control					
Rat No.		9	16	32	46	6	17	47	48		
WEIGHT, g	55th day	158	170	124	108	152	131	133	147		
	85th day	256	263	235	209	253	236	240	248		
	105th day	282	291	280	242	294	262	269	282		
GAIN, g	55-85 days	98	93	111	101	101	105	107	101		
	85-105 days**	26	28	45	33	41	26	29	34		
	55-105 days	124	121	156	134	142	131	136	135		

*Died on Day 90. Evidence of intestinal disturbance.

**Starved on Days 90 and 91, so gain for period is low.

seen that in reality only two of the four surviving rats in this group showed a gain less than the controls. The values that are below normal have been italicized. None of the fat-stimulated group was different from his controls.

Inspection of Table 3 does show one consistent change, and that for both groups. This is the drop in total calories of each experimental group below the value for its control group. This is particularly striking in the case of the fat-consuming group, because the stimulated rats grew as much as the controls.

A 48-hour period of starvation (water only was available) did not alter the patterns of selection. The animals simply consumed more food for a day or two and gained back all of the weight that they had lost. Starvation did not produce any differences in behavior patterns that could be observed.

The greater consumption of salts by the sucrose-control group is not considered significant because of the possibility of spillage which would show up as a relatively large difference for a substance consumed in small quantities. The shift of the fat-control group to sugar has already been mentioned as due to an unstable pattern that could not be avoided. It was pointed out (25) that young rats on self-selection consumed mainly fat for calories, but that the same rats when older ate predominantly sucrose.

4. *Experiment 3 (Type: Long-Passive-Non-Hunger)*

The previous experiment indicated a change in total food consumption, but no shift in patterns of self-selection. The present experiment therefore, is similar except that complete purified diets have been used. This was believed desirable in order to check the changes found in Experiment 2.

a. *Animals.* Twenty weanling rats were divided into groups of five each and fed the purified ration and water *ad libitum*. They were weighed daily at first in order to regroup and match weights. During the experiment on stress they were weighed every three to five days.

b. *Food and water consumption.* The food intakes were measured daily for each of the four groups of five rats. The water consumption of the groups, was, in most cases, measured for a two-day period.

c. *Apparatus.* The stress box was that employed in the previous experiment except that two of the apertures were closed so that the purified diet was accessible at either of two containers in each side of the box. The schedule of stimuli and other details of the procedure are listed chronologically in Table 5.

TABLE 5
SEQUENCE OF PROCEDURES FOR EXPERIMENT 3

Day of age	Treatment
22	Given purified ration and water.
32	Groups matched on basis of weight.
35	Initiated food intake measurement.
44	Placed Groups I and III in stress box to adjust to environment. No stress.
50	Stress started: Stimuli 24 hours a day. Shock.....2 minute interval at 30 volts. Buzzer.....3 minute interval. Air Blast5 minute interval.
54	Shock set at 37 volts, intervals the same.
58	Shock set at 46 volts.
64	Shock at 1 minute intervals at 50 volts. Other stimuli same as above.
68	Shock at 55 volts, intervals the same.
82	Shock at 70 volts.
83	Shock at 0 volts.
84	Shock at 55 volts.
86	Shock at 70 volts. Rats in individual cages discontinued.
88	Shock at 55 volts.
90	Remove animals from box; place controls, groups II and IV, in box with shock at 46 volts and other stimuli off.
91	Shock shut off (0 volts).
95	Shock at 46 volts. No other stimuli.
96	Shock off.
97	Experiment discontinued.

d. Results. The weights at selected intervals of the stress groups (Groups I and III) and the control groups (Groups II and IV) are presented in Table 6. In a similar manner the food and water consumption for five-day periods are summarized in Table 7.

Because of normal variations in food intakes of rats no consistent trend is evident in Table 7. The totals for Groups I to IV for the 61st to 90th day are 85.3, 86.7, 84.7, and 91.3 respectively. These show that there

TABLE 6
GROWTH OF RATS SUBJECTED TO NOXIOUS STIMULI FOR LONG PERIODS

Group	Days of age	30	40	50	60	69	78	86	90
I									
Stress	Weight, g.	64.2	108.8	172.0	195.6	215.2	228.8	248.5	254.5
II									
Control	Weight, g.	67.6	113.6	169.0	209.4	245.6	264.4	285.2	294.2
III									
Stress	Weight, g.	65.2	111.2	152.2	182.8	210.0	226.0	245.8	248.0
IV									
Control	Weight, g.	67.0	112.2	160.4	201.4	233.0	252.6	274.0	285.2

TABLE 7
FOOD AND WATER CONSUMPTION OF RATS SUBJECTED TO NOXIOUS STIMULI FOR LONG PERIODS

Group	Days of age	41-45	46-50	51-55	56-60	61-65	66-70	71-75	76-80	81-85	86-90
I	Food*	15.7	16.5	16.1	15.4	13.8	13.7	13.6	15.6	14.8	13.8
Stress	Water	—	16.4	18.2	22.8	20.3	16.4	17.7	16.8	17.6	16.5
II	Food	14.4	16.1	17.2	14.8	14.7	12.3	14.6	14.8	15.8	14.5
Control	Water	—	20.0	20.5	21.2	21.4	18.3	17.9	18.5	18.3	21.7
III	Food	13.4	15.7	15.0	14.8	13.9	14.6	14.3	15.4	13.7	12.8
Stress	Water	—	16.0	16.8	21.6	14.6	15.6	15.0	15.3	15.7	15.1
IV	Food	14.9	16.3	15.1	14.8	20.4	16.9	17.2	19.6	17.7	16.7
Control	Water	—	17.4	17.4	19.3	19.6	17.4	18.2	20.4	19.1	20.3

*Food in g./day/rat and water in cc./day/rat.

was no change in food intake in the experimental groups. Table 6, however, shows a definite decrease in weight gains of the groups under stress. The t score for the difference between weights of the two groups for Day 60 is 2.49, which is significant just slightly below the 2 per cent level of confidence. For Day 78, t is 3.96, a highly significant difference. It is evident that there is no difference on Day 50 when stress was initiated.

Because of an unusual rise in environmental temperature which resulted in lowered and erratic growth rates and food consumptions, the data are not tabulated beyond Day 90.

5. Experiment 4 (Type: Long-Active-Hunger)

The first experiment involving hunger as the drive was preliminary. The animals were in groups of five in the experimental and in the control series. The methods and results will be presented briefly before proceeding to Experiment 5.

a. *Animals.* Twenty male, weanling rats were placed on a purified diet and water *ad libitum* on their 22nd day of age, and were continued on this regimen for the duration of the experiment.

b. *Food and water consumption.* The purified diet was the same as the one previously used in the other experiments. The food and water intake of each group of animals was measured daily. They were weighed three times per week.

c. *Apparatus.* The stress box used was a modified version of the apparatus described in Experiments 2 and 3. The present model had only one feeding aperture for each of the two compartments of the box. Each aperture had to be approached by a grid ramp. This ramp was electrified each time the animal put his head through the aperture to eat, using as its regulating apparatus the capacitance-controlled electronic relay described elsewhere (24). The voltage applied to the ramp could be varied at the experimenter's discretion. However, because of variability in skin resistance from animal to animal the current affecting the individual animal would vary for any given voltage on the grid.

d. *Procedure.* The 20 animals used in the investigation were divided into four groups of five each, with respect to weight. Two groups were designated stress and were placed one to each of the two compartments within the stress box. The control groups were housed in two wire mesh cages, whose sides were covered with wrapping paper to make conditions similar to the interior of the stress box. Animals were left in their living

quarters for 10 days preceding the introduction of stress to the experimental group.

When the animals were 34 days of age, the stress equipment was activated, subjecting the experimental group to an electric shock each time that they ate and an air blast each time that they drank. The shock was begun at 15v. and was gradually raised until on the 42nd day of the experiment it had reached 37v. with no apparent effect on the animals' behavior. As a result the voltage was more rapidly raised after this to 55, 70, 85 volts. For the remaining days of the experiment the voltage was varied irregularly, depending largely upon the results of the previous day's food consumption. The air blast was discontinued on the 60th day of the experiment. The study was terminated when the animals were 100 days of age.

e. Results. Two general patterns of food consumption emerged as a result of the variation of voltage in this study. In those instances in which the voltage was raised, the food intake of the experimental animals either dropped significantly the day that the voltage was increased, or slacked off gradually over a period of two or three days, during which the raised voltage was continued. When the voltage was dropped a similar set of patterns was evidenced. The food consumption was either increased significantly immediately, or increased steadily over a period of two or three days. Table 8 shows some examples of these responses.

Considering the fact that there were five animals in each compartment of the experimental box, it is difficult to make any generalization concerning these seemingly diverse responses. It is possible that one or two animals in each group affected the data sufficiently to make it appear as it does. Absolute control over the response of the animal is impossible without special equipment in view of the individual differences in skin resistance from animal to animal, and from day to day.

The experimental animals did not grow normally and were significantly below the weight of their controls. This may be attributed to the administration of shock at feeding which discouraged normal eating habits. The weights at 10-day intervals are given in Table 9.

This study suggests that a similar type of stress under circumstances which would allow for the measurement of food consumption of the individual animals would reveal more reliable data. Such a study follows this one. The equipment for a design is more complex, necessitating as it does separate "stress chambers" for each of the animals in order that the individual food and water consumptions may be measured.

TABLE 8
EXAMPLES OF FOOD INTAKE PATTERNS

Day	Volts	Stress I g.	Stress II g.	Con I g.	Con II g.
<i>A. Voltage raised</i>					
<i>I. Gradual decrease in food consumption</i>					
67	70	13.8	15.0	14.2	17.0
68	85	11.4	10.8	13.2	13.6
69	85	15.4	19.0	13.8	17.8
70	85	10.8	14.4	16.2	17.8
71	85	8.2	1.0	15.2	17.2
<i>II. Immediate decrease in food consumption</i>					
60	46	12.8	17.6	17.0	15.2
61	55	10.0	15.0	13.6	15.6
62	70	2.4	1.4	15.2	14.2
<i>B. Voltage lowered</i>					
<i>I. Gradual increase in food consumption</i>					
62	70	2.4	1.4	15.2	14.2
63	55	7.0	14.4	16.0	17.2
64	55	13.2	18.0	15.4	17.0
<i>II. Immediate increase in food consumption</i>					
84	85	3.0	7.4	15.2	17.0
85	70	8.3	9.8	14.4	15.2
86	55	14.0	16.8	16.4	16.6

TABLE 9
WEIGHTS OF ANIMALS IN GRAMS FOR EACH 10 DAYS DURING EXPERIMENT

Days of age	23	33	43	53	63	73	83	93
<i>Experimental Group</i>								
I	42.3	74.0	109.0	133.8	158.4	177.6	190.0	208.3
II	42.2	72.4	111.6	144.6	172.6	192.6	206.0	218.0
<i>Control Group</i>								
III	40.8	76.2	129.8	152.2	192.0	221.2	246.6	256.4
IV	42.6	73.2	128.0	155.9	199.2	226.0	257.2	266.2

6. Experiment 5 (Type: Long-Active-Hunger)

A preliminary experiment was presented (Experiment 4) describing the so-called "long-active" stress situation in which the animal was subjected to punishment each time it ate. This type of experiment has been repeated employing individual stress chambers in order to measure the individual responses of the rats to the situation.

a. Animals. Twenty male, weanling rats were employed. Each rat was housed in an individual wire mesh cage 18 x 18 x 14 inches. Ten of the

rats had free access to food and water at all times. The 10 experimental rats had to approach the food and water sources by means of the ramp in the stress chamber. For the first 18 days these experimental rats had essentially free access to the food because the potentiometer was set at zero. This period was allowed for acclimatization of the animals and for equating them in terms of growth and food consumption.

b. Diet. The ration fed the animals was the same purified diet employed in other experiments.

c. Apparatus. The device for imposing the conflict was a box 6 x 6 inches and 11 inches high inside the large wire mesh cages. In it was a ramp consisting of monel metal bars spaced $\frac{3}{8}$ inch apart to form an electric grid. At the rear of the box (the upper part of the ramp) were two holes, one $\frac{3}{4}$ inch in diameter through which the drinking tube was inserted, and one $1\frac{1}{2}$ inches in diameter behind which was placed the food cup. The alternate bars of the ramp formed an electrified grid. The edge of the larger hole was lined with a metal strip. This metal strip and the water in the drinking bottle were connected to the control grid of the capacitance-controlled electronic relay (24). The one set of alternate bars formed the ground connection for its activation and also was one electrode of the grid for shocking. The relay closed the circuit of the other set of bars (electrode) which was supplied from an adjustable potentiometer. In order to eat or drink the rat had to stand on the grid. When he actually touched the water in the tube or placed his head through the hole behind which the food was placed he actuated the relay and hence received a shock through the paws. Eating behavior was thus punished, producing a conflict between two incompatible responses: one, to obtain nutriment, and the other, to avoid pain.

d. Procedure. After equating the animals and allowing them to attain a weight of 100 grams the potential on the grids was gradually increased until there was a noticeable drop in food consumption. This level of shock was maintained until the rats withstood it and reached a nearly normal rate of food intake. The voltage was then again increased until the animals no longer adapted to a given level of shock.

At this time the potential applied to the grids was decreased to a level to which they had previously become accustomed. It was shown in the preliminary experiment that two types of response could occur. The rat could eat normally as he had done previously at that level of shock or he would eat less because of a new factor entering into the situation. This factor might be a conditioned fear response or a sensitization of the sensory recep-

tors. In either of the latter cases there would be a response governed by the previous and not the present situation. That is, a condition would have been produced in which a given type of behavior endured beyond the stress situation.

When, after one or more decreases in the voltage, the animals had reached a normal food consumption level, the voltage was again increased stepwise until the food and water intake remained low, and the same procedure was re-instituted.

e. Results. The general trend of the results in terms of group average of the experimental and control rats are tabulated for each day in Table 10. At those points where the voltage was decreased and responses indicate the lasting effects of the motivational conflict, the data for the experimental animals are tabulated individually in Table 11.

For the first period, Day 58, all animals immediately resumed eating a normal amount of food. In all subsequent tests of this point it is evident that the results are variable, not only between rats, but for the same rat at different times. Rat No. 12 never showed any prolongation of the effect. Those food intakes showing some after-effect are italicized. Except for two cases, the effect was never marked beyond one day.

D. DISCUSSION AND CONCLUSIONS

The experiments presented in this report have been designed to determine whether rats subjected to motivational stress deviate in their growth rates or in their levels or patterns of food consumption. Previous studies in the field of experimental psychopathology have dealt in part with observations whose interpretation was at times subjective. In addition, the subjects have been animals whose genetic background was unknown or uncontrolled. Objective measurements were often made on the various physiological systems such as respiration rate and pulse. However, although food was an active factor in the experimental situation because the hunger drive was frequently used as motivation, little or no attention was given to the nutritional status of the animal or to the quantity or quality of the diet other than to restrict it sufficiently to produce a high degree of motivation. The approach to the present series of investigations was dictated by a consideration of these various factors.

The rat was the one species that lent itself readily to a systematic attack on the problem. Highly inbred strains are available. Use of such animals eliminates much of the variability arising from genetic differences. Perhaps

TABLE 10

FOOD INTAKE AND WEIGHT OF CONTROL RATS AND OF EXPERIMENTAL RATS SUBJECTED TO SHOCK

Days of age of rats	Experimental		Average weight of rats g.	Conflict from grid volts	Control		Average weight of rats g.
	Food intake g. rat/day	Water intake ml. rat/day			Food intake g. rat/day	Water intake ml. rat/day	
31	—	—	67.6	0	—	—	67.3
35	—	—	75.5	0	—	—	74.8
38	11.3	—		0	13.0	—	
39	11.6	—		0	12.3	—	
40	11.2	—	101.1	0	12.7	—	100.9
41	11.4	—		20	13.7	—	
42	9.1	—	101.8	30	14.2	—	107.2
43	10.0	12.2		30	12.0	17.1	
44	6.6	12.2		40	11.4	17.1	
45	12.4	12.2	113.0	40	16.2	17.1	124.8
46	9.5	12.2		50	15.7	17.1	
47	4.3	12.2		50	9.6	17.1	
48	12.6	12.2		50	14.4	17.1	
49	8.3	12.2		60	12.6	17.1	
50	8.6	12.5	124.6	60	14.9	17.7	148.1
51	7.6	12.5		60	12.3	17.7	
52	11.0	12.5		60	12.4	17.7	
53	8.4	12.5		60	13.7	17.7	
54	6.3	12.5		60	11.7	17.7	
55	13.0	12.5		60	16.9	17.7	
56	8.6	14.8		60	13.6	19.7	
57	9.2	14.8		60	14.7	19.7	
58	12.8	14.8	151.8	50	15.4	19.7	179.3
59	11.2	14.8		50	14.5	19.7	
60	11.6	14.8		40	16.4	19.7	
61	10.6	14.8		30	12.8	19.7	
62	13.9	17.1		0	14.0	20.6	
63	14.5	17.1		30	15.1	20.6	
64	12.4	17.1		30	14.3	20.6	
65	13.7	17.1		40	14.8	20.6	
66	11.0	15.5	175.9	50	14.4	21.9	208.1
67	11.5	15.5		50	13.2	21.9	
68	8.9	15.5		60	14.5	21.9	
69	15.0	15.5		60	14.2	21.9	
70	6.7	15.5		65	13.4	21.9	
71	2.7	12.4		65	15.5	22.7	
72	4.3	12.4		65	14.9	22.7	
73	7.8	12.4	177.9	50	13.0	22.7	232.5
74	13.8	12.4		50	17.2	22.7	
75	11.3	12.4		50	14.6	22.7	

TABLE 10 (Cont'd)

Days of age of rats	Experimental		Average weight of rats g.	Conflict from grid volts	Control		Average weight of rats g.
	Food intake g. rat/day	Water intake ml. rat/day			Food intake g. rat/day	Water intake ml. rat/day	
76	11.9	14.5	190.9	60	14.3	21.9	241.5
77	4.9	14.5		65	12.9	21.9	
78	12.3	14.5		65	18.4	21.9	
79	12.9	14.5		70	17.3	21.9	
80	9.2	14.5		75	13.8	21.9	
81	8.9	11.6	192.3	80	15.9	23.4	258.4
82	5.6	11.6		85	12.8	23.4	
83	10.5	11.6		85	15.1	23.4	
84	5.3	11.6		90	14.2	23.4	
85	12.1	11.6		70	14.1	23.4	
86	7.1	10.1	201.2	70	15.3	25.3	278.4
87	11.7	10.1		70	16.1	25.3	
88	9.8	10.1		80	13.1	25.3	
89	5.6	10.1		90	14.7	25.3	
90	5.7	10.1		90	15.6	25.3	
91	8.7	13.3	212.6	80	13.4	25.9	293.3
92	9.2	13.3		80	13.3	25.9	
93	10.4	13.3		80	14.6	25.9	
94	11.6	13.3		70	14.5	25.9	
95	14.0	11.9		70	16.6	27.3	
96	8.5	11.9	229.2	80	12.6	27.3	300.8
97	10.4	11.9		80	14.9	27.3	
98	8.8	11.0		90	15.3	25.7	
99	6.9	11.0		95	15.1	25.7	
100	9.2	11.0		95	17.4	25.7	
101	9.2	11.0	229.2	90	14.7	25.7	300.8
102	7.2	11.7		95	16.1	26.2	
103	5.2	11.7		95	15.2	26.2	
104	6.5	11.7		85	15.1	26.2	
105	7.0	11.7		85	11.3	26.2	
106	9.1	11.7	229.2	75	—	—	300.8
107				0			

even more important is the fact that the nutritional requirements of the rat have been well defined for optimum growth. By the use of purified diets, one is in a position to know the nutritional status of the animals and to vary it at will.

The methods employed in an attempt to induce motivational stress were geared to the repertoire of responses available to the rat and yet could

probably be applied, with some modification, to other species higher in the phylogenetic level. Furthermore, in consideration of the possible rôle of nutrition, part of the experiments were designed to avoid the use of starvation

TABLE 11
FOOD CONSUMPTION EXPRESSED AS GRAMS PER DAY OF INDIVIDUAL EXPERIMENTAL RATS
AT SELECTED PERIODS WHEN VOLTAGE WAS DECREASED

Day	Volts	Rat number									
		12	13	15	21	24	27	28	30	34	37
56	60	4	9	7	8	8	10	12	9	10	9
57	60	13	8	9	9	9	9	8	9	12	6
58	50	14	13	12	15	11	11	12	10	11	19
59	50	12	10	9	11	9	9	10	9	19	14
60	40	10	20(a)	12	13	14	9	10	12	13	24(a)
61	30	12	9	10	10	10	12	9	10	12	4
62	0	13	17	17	15	13	13	12	12	13	5(b)
70	65	6	5	6	3	4	12	8	7	8	8
71	65	3	3	8	3	0	2	3	3	0	2
72	65	10	8	10	8	2	1	3	0	0	1
73	50	13	13	10	11	6	8	4	2	4	7
74	50	14	18	18	12	16	14	9	11	9	17
83	85	16	16(c)	5	9	12	9	9	3	13	13
84	90	6	1	7	6	5	4	9	0	8	7
85	70	14	8	15	16	14	8	9	—(d)	13	12
86	70	10	2	11	7	9	5	3	5	8	12
87	70	13	8	12	12	16	11	9	10	13	13
89	90	8	4	4	4	5	3	7	9	6	6
90	90	10	2	2	10	8	0	6	6	7	6
91	80	14	5	4	10	10	4	11	11	8	10
92	80	11	8	9	9	6	9	9	10	10	12
93	80	16	8	6	11	11	9	9	12	11	11
94	70	13	14	8	13	9	8	12	11	14	14
95	70	13	16	15	11	17	13	13	14	17	11
99	95	7	3	10	11	4	8	2	10	7	(e)
100	95	10	9	10	10	6	12	5	14	7	
101	90	11	11	11	11	10	9	4	8	8	
102	95	9	3	9	13	13	7	0	12	6	
103	95	8	4	9	2	6	2	13(f)	7	9	
104	85	12	5	10	7	12	4	2	7	12	
105	85	15	6	9	0	10	7	11(f)	7	9	
106	75	15	8	13	12	13	10	1	8	11	

(a) Rat pushed chamber away; had free access to food.

(b) Food cup turned around in holder; limited rat's access to food.

(c) Shock turned off for this rat because it ate no food on the two previous days.

(d) Food spilled; probably only a small amount eaten for rat lost eight grams in two days.

(e) Rat died; apparently electrocuted.

(f) Rat's foot burned and swollen from shock; shock turned off for these two days. Large amount eaten probably accounts for decreased intake on each of the following days with less motivation to withstand shock in order to eat.

in order to motivate the animals. The experiments were also graded to subject the animals to smaller and larger amounts of stress in terms of its duration each day.

Although the animals were observed for evidence of behavioral changes, the major criteria were the growth rates and the level of food consumption. In discussing the individual experiments these will be the factors given the greatest weight.

The first experiment was an extension of the work of Maier and his co-workers (17, 18). Because of the emphasis which he placed on this type of conflict it seemed quite possible that it might yield evidence of changes in the rats. The results demonstrated rather clearly that such a short-term stress each day had no effect on the food intake or the weight gain of the animals. The subsequent experiments were then set up in an effort to induce stress for 12 to 24 hours each day for periods of 30 days or more.

Experiments 2 and 3 may profitably be considered together. The only differences in the design of the two problems lay in the determination of the patterns of selection in the first of them. Since no changes were noted in the self-selection patterns, the second of these experiments was simplified by employing the complete purified ration. The first of the experiments gave an indication of a decrease in the total caloric consumption by the rats. The results of the third experiment did not confirm the changes in food consumption. The difference may be the result of the two types of diet. Further studies might clarify this point. As pointed out in the results of Experiment 3, there was no change in the food intake during the period of stress. There was, however, a drop in the rate of growth of the rats subjected to the nociceptive stimuli. Examination of the data for individual days indicated that the change was brought about by the electric shock. Thus, when the voltage was increased there was a decrease in weight gain and food consumption for a day or two. It is apparent from this study that painful stimuli produce reactions in the animals which result in a decrease in growth. Since the total food consumption was not altered, the rats evidently consumed sufficient ration at other times to compensate for the temporary decrease.

The conclusion from the above results is that the rats gained less weight on the same caloric intake as the controls. There are two possible explanations of this fact. (a) The stress caused additional activity which consumed more energy. (b) The painful stimuli resulted in less efficient food utilization. Further studies involving measurement of additional variables such as activity and an even closer control over temperature and humidity of the

environment would be necessary to define more clearly the existing relationships in this problem. However, it would seem that the type of stress involved in these experiments is physiological in the sense that overt behavior abnormalities are not present. That is, the changes are those that result from such factors as forced exercise or adaptation to changes in environmental temperatures, etc., and are not those of "experimental neurosis" or behavior abnormalities in the psychological sense. This point of view, however, can only be tested by further investigations.

The interpretation of the results of Experiments 4 and 5 may be approached from two points of view. The decrease in food consumption at certain levels of shock is only a measure of the effectiveness of a given current (in terms of voltage, since the skin resistance changes) in opposing the hunger drive. The subsequent increase in food intake at the same voltage may represent an adaptation of the animal to withstand the shock or an increase in the motivation with continued low levels of food consumption. The lower weight gain is of course a necessary concomitant of the lower caloric consumption. Although these results may be of use in some types of experiments they do not bear directly on the present series of investigations.

The principal objective of the experiments of the "long-active-hunger" type was to determine whether the conflict between food-taking and pain-avoidance would manifest itself by behavior that was no longer suited to the situation. By reducing the effective shock that had originally prevented the animals from obtaining a normal amount of food but relying on the position of the food in a formerly conflictful situation and the presence of a lessened degree of shock that had formerly been too strong for the rat to withstand, one could conceivably find out whether the animal would still avoid the food and thus indicate that the conflict was still operating even though the conditions no longer demanded it. The results showed that the rats quickly adjusted to the situation. Even though the conflictful situation existed for periods of five to 14 days at a time, when the voltage was decreased the rats resumed eating at the normal level within one day. Thus no really lasting effect was produced. A different type of study relating food intake to conflict has recently been reported (27).

The experiments gave little indication of changes in the rate of weight gain and the amount of food consumed by rats in a controlled environment designed to subject them to stress of a conflictful or frustrating nature. Under the conditions of these studies the rat readily readjusted its behavior patterns which are relatively simple. As soon as the situation which interfered

with the animal's ability to obtain food and maintain normal metabolic functions for growth ceased to exist, the rat adopted the former mode of activity or behavior.

It should be realized that these studies are essentially preliminary or screening tests using relatively small groups of animals.

The basic systematic approach to the problem of the rôle of stress in growth and food consumption, however, would seem to be applicable to other species and perhaps to the rat under somewhat different conditions. It is the opinion of the authors that the interrelationships between nutrition or metabolism and behavior, both normal and abnormal, is an important and fruitful field for further investigations.

E. SUMMARY

A series of investigations were designed to determine whether rats could be subjected to forms of motivational stress that would result in a change in their growth rates or levels of food consumption. These criteria were chosen as objective measures of altered behavior.

The experiments were classified on the basis of the variables involved. The three factors in the studies were the duration of the stress in terms of time each day, the rôle of food in the situation, whether present or not, and the rôle of the animal which either was active in initiating the conflict situation or was subjected to it passively. Three combinations of these variables were employed in five experiments.

An active conflict situation for short periods each day did not alter the growth rate or amount of food consumed in the living quarters. Rats subjected to long periods of noxious stimuli exhibited a decrease in growth rate which may have resulted from one of several factors. Further experiments are necessary to define more clearly the relationships involved. A conflict between food-getting and shock avoidance decreased the food consumption during the conflict but had no lasting effects in that food intake became normal shortly after a decrease in the degree of the conflict.

The studies gave little evidence that rats show any marked deviations in behavior that result in alterations in food consumption or growth. However, there is some indication that modification of the basic approach might yield positive results with the rat. Other evidence in the literature on "experimental neuroses" suggests that these methods may be more effective in species higher in the phylogenetic level.

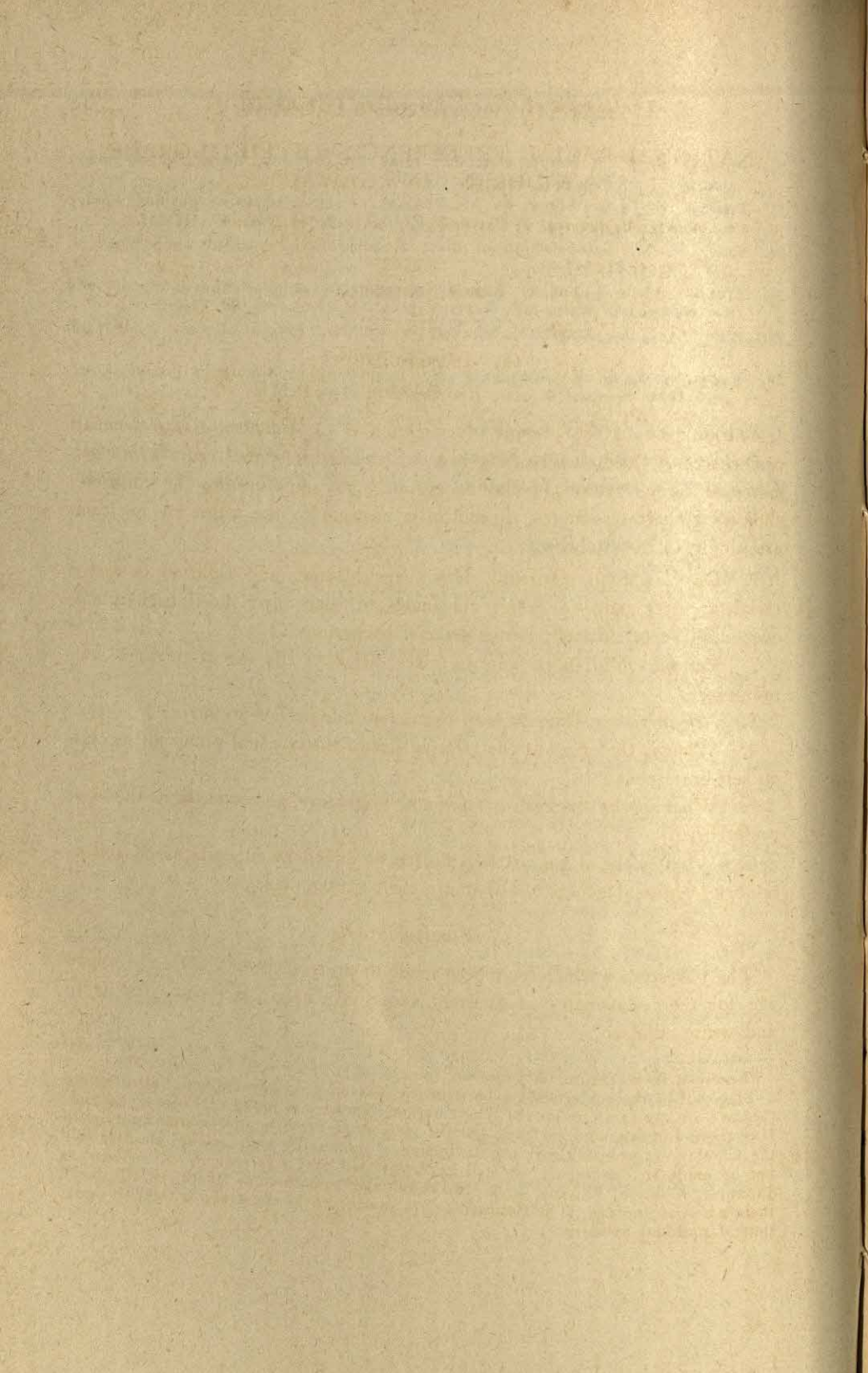
REFERENCES

1. ANDERSON, O. D., & PARMENTER, R. A long-term study of the experimental neurosis in the sheep and dog. *Psychosom. Med. Monog.*, 1941, 2, Nos. 3 & 4.
2. CANNON, W. B. Bodily Changes in Pain, Hunger, Fear, and Rage. New York: Appleton Century, 1939.
3. ———. Hunger and thirst. In MURCHISON, C., (Ed.): *The Foundations of Experimental Psychology*. Worcester, Mass.: Clark Univ. Press, 1929. (Pp. 434-448.)
4. COOK, S. W. A survey of methods used to produce "experimental neurosis." *Am. J. Psychiat.*, 1939, 95, 1259-1276.
5. ———. The production of "experimental neurosis" in the white rat. *Psychosom. Med.*, 1939, 1, 293-308.
6. DANIEL, W. J. An experimental note on the O'Kelly-Steckle reaction. *J. Comp. Psychol.*, 1943, 35, 267-268.
7. FINGER, F. W. Experimental behavior disorders in the rat. In HUNT, J. McV., (Ed.), *Personality and the Behavior Disorders*. New York: Ronald Press, 1944. (Pp. 413-430.)
8. ———. Convulsive behavior in the rat. *Psychol. Bull.*, 1947, 44, 201-248.
9. GANTT, W. H. *Experimental Basis for Neurotic Behavior*. New York: Hoeber, 1944.
10. ———. The origin and development of nervous disturbances experimentally produced. *Am. J. Psychiat.*, 1942, 98, 475.
11. HEBB, D. O. Spontaneous neurosis in chimpanzees: Theoretical relations with clinical and experimental phenomena. *Psychosom. Med.*, 1947, 9, 3-16.
12. HUMPHREY, G., & MARCUSE, F. New methods of obtaining neurotic behavior in rats. *Am. J. Psychol.*, 1939, 52, 616-619.
13. HUNT, J. McV., & SCHLOSBERG, H. Behavior of rats in continuous conflict. *J. Comp. & Physiol. Psychol.*, 1950, 43, 351-357.
14. JONES, J. H., & FOSTER, C. A salt mixture for use with basal diets either low or high in phosphorus. *J. Nutr.*, 1942, 24, 245-256.
15. LASHLEY, K. S. The mechanism of vision: I. A method for rapid analysis of pattern vision in the rat. *J. Genet. Psychol.*, 1930, 37, 453-460.
16. LIDDELL, H. S. Conditioned reflex method and experimental neuroses. In HUNT, J. McV., (Ed.), *Personality and the Behavior Disorders*. New York: Ronald Press, 1944. (Pp. 389-412.)
17. MAIER, N. R. F. *Studies of Abnormal Behavior in the Rat*. New York: Harper, 1939.
18. MAIER, N. R. F., GLASER, N. M., & KLEE, J. B. Studies of abnormal behavior in the rat: III. The development of behavior fixations through frustration. *J. Exp. Psychol.*, 1940, 26, 521-546.
19. MASSERMAN, J. H. *Behavior and Neuroses*. Chicago: Univ. Chicago Press, 1943.
20. MOWRER, O. H. An experimental analogue of "regression" with incidental observations on "reaction-formation." *J. Abn. & Soc. Psychol.*, 1940, 35, 56-87.
21. O'KELLY, L. I., & STECKLE, L. C. A note on long enduring emotional responses in the rat. *J. of Psychol.*, 1939, 8, 125-131.
22. PAVLOV, I. P. *Conditioned Reflexes and Psychiatry* (Trans. by W. H. Gantt). New York: International Publishers, 1941.
23. ———. *Conditioned Reflexes* (Trans. by G. V. Anrep). London: Oxford Univ. Press, 1927.

24. PILGRIM, F. J. A simple electronic relay for counting, timing, or automatic control. *J. of Psychol.*, 1948, **26**, 537-540.
25. PILGRIM, F. J., & PATTON, R. A. Patterns of self-selection of purified dietary components by the rat. *J. Comp. & Physiol. Psychol.*, 1947, **40**, 343-348.
26. SCOTT, E. M. Self-selection of diet: I. Selection of purified components. *J. Nutr.*, 1946, **31**, 397-406.
27. TOLCOTT, M. A. Conflict: A study of some interactions between appetite and aversion in the white rat. *Genet. Psychol. Monog.*, 1948, **38**, 83-142.
28. WITKIN, H. A. Abnormal behavior in animals. *Psychol. League J.*, 1939, **3**, 75-83.
29. YACORZYNSKI, G. K. Necessary and facilitating conditions in producing experimental neurosis in cats. *Am. Psychol.*, 1946, **1**, 245.

*Quartermaster Food & Container
Institute for the Armed Forces
Chicago Quartermaster Depot,
U. S. Army
1819 West Pershing Road
Chicago 9, Illinois*

*Department of Psychology
University of Pittsburgh
Pittsburgh, Pennsylvania*



NATIONAL-RACIAL PREFERENCES OF FIFTH-GRADE CHILDREN IN HAWAII*

Department of Psychology, University of Hawaii

DORIS SPRINGER

A. INTRODUCTION

1. *Problem*

The purpose of the present investigation is to study the national-racial¹ preferences of children who have been in schools for several years with classmates of varied national-racial backgrounds, by determining the children's choices of their classmates. Specifically, answers to the following questions are sought in the study:

1. Do Caucasian, Oriental, Hawaiian children, and children of mixed racial ancestry tend to prefer classmates of their own racial backgrounds over children of different national-racial backgrounds?
2. Do the children of various races differ in the extent of their self-preference?
3. Are there sex differences in the magnitude of self-preference?
4. What is the effect of the relative size of the national group in the class on self-preference?
5. What is the strength of national preference as compared with racial preference?
6. What is the status of the children of mixed racial background and of minority national groups? What are their preferences?

2. *Previous Studies*

The few studies which have been made of preferences of children of school age for their classmates of different races have been concerned with negro and white children.

*Received in the Editorial Office on July 24, 1951.

¹Since all of the subjects were born in the Territory or in mainland United States, they are all American nationals. Their national ancestry is Japanese, Chinese, Filipino, Hawaiian, English, and so forth. In the interest of brevity, and in conformity with the popular usage in Hawaii, the distinction of nationality and national ancestry will not be made in this report. It will be understood that reference to the subjects as Japanese, Chinese, Filipino, and Hawaiian, and combinations of these will mean their national ancestry. The Caucasians are considered as one group without designation of national ancestry.

Koch (2, 3) studied the development of racial, nationality, and skin pigmentation preferences in children of school age by comparing the responses of white and negro, and of Italian and non-Italian children. The paired comparisons method was used. Koch found that the white children preferred their white classmates from the earliest to the highest grades studied. On the average, the proportion of choices made by Negro pupils which favored negroes increased steadily, and from the eighth grade on they uniformly favored their negro classmates. The white girls showed, on the average, a stronger preference for their white classmates than did boys. The white boys studied tended to show their preference for their own race more when judging girls than boys, whereas the reverse seemed to be true in the case of negro boys. When members of a given race were in a small minority, they tended to express a stronger preference for members of their own race who were their own sex, while the reverse obtained in the case of members of the opposite sex. There was no statistically significant nationality cleavage discovered in the Italian and non-Italian children. Skin-color cleavages were evident, with each color group tending to prefer itself.

Moreno (4), using the sociometric technique, made a study of social structures evolved by groups of negro and white children in the public schools in New York. Each child was asked to choose two others beside whom he would like to sit and to give the reasons for his choices. Nationality and race groupings became noticeable in the choices by grade four, but were not mentioned in the interviews as reasons for their choices until grade five.

Criswell (1) used the sociometric technique in which each child was asked to make a first and second choice of children beside whom he would like to sit. He concluded that although self-preference may exist in negro and white children even in the kindergarten, the "fifth grade level stands out as the point at which mutual withdrawal of the races crystallizes as the characteristic group pattern." White children were found to surpass negroes in self-preference. An attempt was made to analyze the effect of the size of the minority race relative to the majority race in the various classes. In girls' groups an increase in the minority size produced increased rejection of the minority by the majority, while an opposite effect was produced in the case of the boys. The minority did not exhibit a consistent change of attitude toward the majority with an increase in size. White majorities exhibited more self-preference than did similarly constituted negro groups.

In the design of the present study, an attempt has been made to obtain data on the relationship of the Oriental and Caucasian children and on the

relationship of the two nationalities, Chinese and Japanese, which might serve to broaden or limit the application of the conclusions of the previous studies on race and nationality preferences. In addition, the attitude of their classmates toward children of mixed racial background, including Hawaiians, and the preferences of the latter are considered.

B. SAMPLE AND PROCEDURE

1. *Subjects*

The subjects were all the fifth grade children in three schools in Honolulu, Hawaii. The schools were selected because they have heterogeneous populations with a sufficient number of both Caucasians and Orientals to permit comparisons between the two races. The school staffs are also heterogeneous with respect to nationality and race. The children have been required to pass English examinations upon entrance to the schools. They live in various sections of the city, and the majority are from sections in which the populations are mixed. All children had been in their classes at least three months.

A single grade was chosen for the study in order to control the age factor. The fifth grade was selected because a group-test method could be used effectively with this age level, and the effect of the nationality of the interviewer would not have to be decided. Also, sex distance in fifth grade children has been found to be great (2) which was desired in the study in making certain comparisons between the two sexes. By the time he reaches the fifth grade, the child in Hawaii has had contacts with children of nationalities and races other than his own for several years.

The distribution of the nationalities enrolled in the classes used in the three schools, combined, is shown in Table 1. The distributions were combined since the separate distributions were not significantly different. (The probabilities that the differences were attributable to chance, based on the chi-square test, were .90 for boys, .20 for girls, and .50 for the total). Caucasians are treated as one group without reference to national ancestry since such distinction is not made by the children themselves (except for Portuguese). "Part-Hawaiian" is the official classification in the census and in popular usage for persons who have any Hawaiian ancestry. The Part-Hawaiians in the study were Hawaiian-Caucasians, Hawaiian-Chinese, and Hawaiian-Caucasian-Chinese.

Of the 254 students (104 boys and 150 girls) enrolled in the classes, responses of 233 (92 boys and 141 girls) were available for the final tabulations due to absences and inaccurate marking. The national-racial

TABLE 1
DISTRIBUTION OF ENROLLMENT OF CLASSES STUDIED, ACCORDING TO NATIONAL-RACIAL
BACKGROUND AND SEX

National- racial classification	Per cent of enrollment		
	Boys (N = 104)	Girls (N = 150)	Total (N = 254)
Caucasian	34%	27%	29.5%
Chinese	19	21	20
Japanese	29	35	33
Part-Hawaiian	11	14	13
Korean	2	.7	1
Japanese-Korean	1	1	1
Japanese-Caucasian	2	0	.8
Japanese-Chinese	1	0	.4
Chinese-Caucasian	0	.7	.4
Filipino	0	.7	.4
Negro	1	0	.4

distribution of the subjects was virtually the same as that of the class enrollment, with no deviations as large as one per cent.

2. Procedure

The method of paired comparisons was used. Each child was asked to express his preference when the name of each classmate was paired with the name of every other classmate of the same sex and of a different national ancestry. A mimeographed list of the pairs of names was given to the child with the instructions to underline the name in each pair which he liked the better of the two pupils and to omit the pairs containing his own name. No time limit was imposed, and the time required ranged from five to 20 minutes for the actual marking of the lists.

Provisions were made in preparing the lists to minimize the effect of position of the name on the choice. The order of each child's name in the pair was random (i.e., occurred first as often as last). The position of the pairs in the list was determined by random numbers. The sheets of names were arranged in two different orders for each class.

3. Statistical Analysis

The results are reported separately for boys' choices of boys, girls' choices of girls, boys' choices of girls, and girls' choices of boys, and for the different national and racial groups. Usually the data are in the form of percentages of total choices made by the subjects which favored the national-racial group to which the subject belonged. The total number of choices made was 40,231, an average of 173 per pupil. In some analyses the percentage of the

individual's choices of his own national or racial group when paired with other groups has been used as a "percentage-preference" score (3).

Differences between percentages are considered real if they are significant at the one per cent level (i.e., if $P = .01$ or less, or if the difference is 2.58 or more times the standard error of the difference). The same level of confidence was used to determine whether a percentage differs significantly from the percentage expected by chance (i.e., from 50 per cent).

The emphasis is on the responses of the three major groups of subjects, namely, Caucasian, Chinese, and Japanese, when they made choices between the children of these three groups. The responses of the remaining subjects (most of whom have mixed racial ancestry) are reported separately and in less detail, as are the choices of the major groups of subjects when considering pairs in which this mixed group is involved.

C. DISCUSSION OF RESULTS

1. *Racial Preferences of Caucasian, Chinese, and Japanese Children*

The percentages of the total number of choices made by Caucasian, Chinese, and Japanese subjects which favored classmates of their respective national-racial backgrounds in pairs involving the two races are shown in Table 2. When the Japanese and Chinese names are combined to make possible a comparison of the choices for the two races, Oriental and Caucasian (Column 3), in every instance, except one, the judges showed a favorable bias for the race to which they belonged. There is evidence, however, that different national-racial groups varied in the degree of their racial preference.

The Chinese and Japanese girls preferred Oriental girls when paired with Caucasian girls by a significantly larger percentage than that by which Caucasian girls preferred girls of their own race in selecting from the same pairs. When boys made choices of classmates of their own sex, the Japanese showed a greater preference for Orientals in the Oriental-Caucasian pairs than did Caucasian boys in judging the same pairs. The difference for choices of boys by boys in the case of Caucasians and Chinese was not clearly significant and was the only comparison of the two races in which the greater race preference was shown by Caucasians.

When the children being judged were not the same sex as the judges, the Chinese and Japanese showed a reliably stronger preference for Orientals than Caucasians did for members of their racial groups.

Children of both races revealed a preference for their own race, but the general trend was for Oriental children to show a significantly greater racial self-preference than was found in the Caucasian children.

TABLE 2

PREFERENCES OF CAUCASIAN, CHINESE, AND JAPANESE CHILDREN FOR CLASSMATES OF THEIR RESPECTIVE NATIONAL-RACIAL GROUPS

National-racial background of subject	Per cent of choices of own national-racial group		
	Caucasian-Chinese Pairs	Japanese-Caucasian Pairs	Oriental-Caucasian Pairs
<i>Boys' choices of boys</i>			
Caucasian	70%	56%	61
Chinese	40	63	54*
Japanese	60	76	70
<i>Girls' choices of girls</i>			
Caucasian	64	52*	58
Chinese	76	57	66
Japanese	66	77	72
<i>Boys' choices of girls</i>			
Caucasian	55	55	55
Chinese	76	66	70
Japanese	56	64	61
<i>Girls' choices of boys</i>			
Caucasian	54	37	44
Chinese	53*	61	58
Japanese	41	75	62

*Per cents not significantly different from 50 per cent, or chance.

2. *Magnitude of Racial Self-Preference for the Two Sexes*

Another analysis which can be made from Table 2 is a comparison of the degree of racial self-preference for the two sexes when choices were made from pairs of children of the same sex as the judge. The Caucasian boys and girls chose Caucasians of their respective sexes to approximately the same extent. Likewise, differences in the percentages of choices of own national or racial group by Japanese boys and girls were not significant. In one instance the difference between choices of boys and girls was significant for the Chinese subjects. The Chinese boys chose fewer boys of their own national group in the Caucasian-Chinese pairs than did the Chinese girls when they selected girls in the Caucasian-Chinese pairs. However, the sex difference was not found when the Chinese boys and girls were choosing from Caucasian-Japanese pairs.

The conclusion appears to be that boys and girls in the study did not differ in their degree of national or racial self-preference when choosing classmates of their respective sexes.

3. *Relation of Sex of Judge and Sex of Children in Pairs Being Judged*

Studies of negro and white children have suggested that differences occur in extent of racial preference when the children being judged are the same sex as the judges, and when they are members of the opposite sex, although the trends reported are not consistent for both races (3). The data in Table 2 permit such a comparison of magnitude of the racial self-preferences of boys when they chose boys and girls and of girls when they judged classmates of the two sexes. With the exception of Chinese boys, the children showed a stronger bias toward their own national-racial group when selecting members of their own sex than when selecting members of the opposite sex.

From the same table an analysis can be made of the racial self-preferences of the two sexes when they judged the same pairs of names (e.g., boys' and girls' choices of boys). When indicating their preferences in the male Caucasian-Chinese pairs, Caucasian boys chose significantly more Caucasians than did the Caucasian girls. In the case of the Chinese subjects, the trend was reversed, i.e., the girls chose significantly more children of their own nationality in the Chinese-Caucasian male pairs than did the Chinese boys. The Japanese boys and girls each favored the Japanese boys over the Caucasian boys in approximately three-fourths of their choices.

When selections were made from pairs of girls, preference for members of their own racial groups is not significantly different for Caucasian boys and girls. The Chinese boys and girls favored the Chinese girl in the Caucasian-Chinese pairs to exactly the same extent (approximately three-fourths of the time). However, in the Caucasian-Japanese pairs of girls, the latter were chosen significantly more often by Japanese girls than by Japanese boys (77 per cent and 64 per cent, respectively).

When the pairs of names were judged by children of the same sex as those in the pairs and by children of the opposite sex, differences in the extent of national or racial preference occurred in about half of the comparisons. The stronger racial preference was generally found where the pairs of classmates were the same sex as the judge. It would seem that factors other than nationality and race which influence selections may, and did in many instances, become relatively more important when making selections of the opposite sex than when making selections of own-sex pairs. It may be that since the fifth grade child has less contact with members of the opposite sex than he has with members of his own sex group, more of his choices of the former are random, thus minimizing the bias for his own national or racial group.

4. Comparison of Racial-Preference and Nationality-Preference for Chinese and Japanese Subjects

The relative strength of national and racial self-preferences can be compared for Chinese and Japanese subjects since these two nationalities were paired with each other as well as with Caucasians. From Table 3 it may be noted

TABLE 3
COMPARISON OF RACIAL SELF-PREFERENCE AND NATIONALITY SELF-PREFERENCE OF CHINESE AND JAPANESE CHILDREN

Nationality of subject	Per cent of choices of own nationality when paired with:		
	Caucasian	Chinese	Japanese
<i>Boys' choices of boys</i>			
Chinese	40%	—	46%*
Japanese	76	60	—
<i>Girls' choices of girls</i>			
Chinese	76	—	45*
Japanese	77	71	—
<i>Boys' choices of girls</i>			
Chinese	76	—	50*
Japanese	64	62	—
<i>Girls' choices of boys</i>			
Chinese	53*	—	55*
Japanese	75	70	—

*Per cents not significantly different from 50 per cent, or chance.

that in no instance did the Chinese boys or girls choose a significantly greater percentage of Chinese in the Chinese-Japanese pairs than would be expected by chance (50%). On the other hand, in every instance Japanese pupils chose Japanese classmates significantly more than 50 per cent of the time when judging the same pairs.

When the percentage of Chinese chosen in the Chinese-Japanese pairs is compared with the percentage chosen in the Caucasian-Chinese pairs, there is no statistically significant difference when the pairs were boys. When the pairs were girls, however, subjects of both sexes chose a greater proportion of their own nationality when paired with the other race than when paired with the other nationality. When the Japanese subjects selected both boys and girls, the magnitude of their national self-preference was greater when matched with the Caucasians than when matched with the Chinese. All differences were statistically significant except for boys' choices of girls.

In general, the Oriental subjects showed greater self-preference when their nationality was paired with another race (Caucasians) than when their

nationality was paired with another nationality of the Oriental race (Chinese and Japanese).

5. *Variability of the Magnitude of Racial Self-Preference*

So far, the responses of the children have been considered in terms of the percentage of the total number of responses in each national-racial group which favored children belonging to their particular national-racial group. Obviously there were variations in the magnitude of racial self-preference of different children in any given national-racial group. To show this variation the percentage of times the subject selected the members of his own race, when choosing between pairs of classmates of two races, was determined. These percentage-preference scores for the three major groups of subjects are presented in two ways: first, in terms of the quartiles; and second, in terms of the proportions of subjects in each group who chose classmates of their own race more than half of the time (percentage-preference scores greater than 50).

If the spread in the magnitude of racial self-preference for the middle 50 per cent of the subjects (the interquartile range) in each group (Table 4) is taken as the measure of variability, then no consistent differences are evident in the variability of the degree of racial self-preference when compari-

TABLE 4
QUARTILES OF DISTRIBUTIONS OF RACIAL PREFERENCE-PERCENTAGE SCORES OF CAUCASIAN, CHINESE, AND JAPANESE CHILDREN

National-racial group of subject	Per cent of choices of own race		
	First Quartile	Second Quartile	Third Quartile
		<i>Boys' choices of boys</i>	
Caucasian	44%	62%	76%
Chinese	44	67	83
Japanese	61	75	92
		<i>Girls' choices of girls</i>	
Caucasian	42	64	74
Chinese	45	57	72
Japanese	57	71	89
		<i>Boys' choices of girls</i>	
Caucasian	45	57	68
Chinese	54	76	82
Japanese	39	54	82
		<i>Girls' choices of boys</i>	
Caucasian	33	50	57
Chinese	42	64	85
Japanese	50	64	86

sons are made (*a*) between the two sexes, (*b*) between the three national-racial groups, and (*c*) between choices for pairs of the judge's own sex and for pairs of the opposite sex. The middle 50 per cent of the subjects varied in the extent of racial self-preference from 23 to 43 percentage points.

Similar observations concerning the magnitude of racial self-preference can be made from the data in Table 4 to those made in Table 2. On the average, the boys and girls preferred their classmates of their own race to about the same extent; the preferences were more marked in Orientals than in Caucasians; and usually racial self-preference was greater when the classmates judged were of the subject's own sex rather than of the opposite sex.

The data in Table 5 present in another form the information on the extent and variability of the racial self-preference of the individual subjects. When

TABLE 5
PERCENTAGES OF CHILDREN WHOSE CHOICES OF CLASSMATES OF THEIR OWN RACE
WHEN MATCHED WITH ANOTHER RACE EXCEEDED FIFTY PER CENT

Nationality of subject	Per cent choosing own race more than half the time			
	Boys' choices of boys	Girls' choices of girls	Boys' choices of girls	Girls' choices of boys
Caucasian	67%	70%	66%	43%
Chinese	68	67	68	60
Japanese	89	80	54	70

choosing members of their own sex, slightly more than two-thirds of the Caucasian and Chinese boys and girls and more than four-fifths of the Japanese boys and girls chose classmates of their respective races in more than half of their selections.

6. *Effect on Racial Self-Preference of Relative Size of the Subject's National Group in the Class*

Various methods were tried in an attempt to determine the effect on self-preference of the subjects when the proportionate size of the national groups by sex varied in the eight classes studied. The problem was complicated by the fact that children of the Oriental race outnumbered the Caucasians in every group. Also, in only one instance (Japanese girls in one class) did the subjects of a given nationality exceed 50 per cent of the total class enrollment of either sex. Consequently, when the subjects were divided according to nationality and sex, the numbers in each category were very small. In spite of the acknowledged limitations, Table 6 is presented here because of the consistent trend found in the data.

TABLE 6
EFFECT OF THE RELATIVE SIZE OF THE CHILD'S NATIONAL GROUP IN THE CLASS ON
SELF-PREFERENCE

Nationality of subject	Median of percentages of preference for own nationality	
	Relative size of national group	
	Small	Large
	<i>Boys' choices of boys</i>	
Caucasian	64%	78%
Chinese	29	89
Japanese	70	89
	<i>Girls' choices of girls</i>	
Caucasian	40	61
Chinese	67	78
Japanese	59	88

The proportions of Caucasian, Chinese, and Japanese boys and girls in each of the eight classes were determined, and the two or three classes with the smallest proportions of each nationality for each sex, and the two or three classes with the largest proportions of each nationality for each sex were selected for the comparison. For each subject in these categories, the percentage of his choices of classmates of his nationality when they were paired with classmates of the other race was determined. The figures reported in Table 6 are the medians of these preference-percentage scores when the subject judged members of his own sex.

For Caucasians, Chinese, and Japanese of both sexes, the tendency to choose classmates of their respective nationalities increased, on the average, as the relative size of the subject's national group increased in the class. The same trend was found when racial preference, instead of national preference, was computed for the Chinese and Japanese subjects.

It may be of interest here, also, that the only cases in which a subject failed to choose any of his classmates of his national group and sex occurred in classes in which his group was small. These subjects were three Caucasian boys, two Chinese girls, and one Chinese boy. On the other hand, only three (all Chinese) out of the 17 subjects (all Orientals) who chose children of their respective national groups 100 per cent of the time when paired with Caucasians, were in classes in which their national group was relatively small. The remaining 14 subjects constituted about one-third of the Japanese and Chinese subjects in the classes in which the size of their respective national groups have been designated as "large" in the table.

Thus, it appears that when the class is made up of different nationality and

racial groups, members of a relatively small national group may vary in the degree of self-preference from zero to 100 per cent, with the former somewhat more likely to occur than the latter. The members of a relatively large national group in a class show a higher average national self-preference than they do when their national group is relatively small, and they are much more likely to prefer representatives of their nationality exclusively in the former classes.

7. Choices of the Minority

The preferences expressed by the subjects in the class other than the Caucasians, Chinese, and Japanese are shown in Table 7. These subjects

TABLE 7
PREFERENCES OF PART-HAWAIIAN CHILDREN AND CHILDREN OF OTHER MINORITY NATIONALITIES FOR THEIR CLASSMATES

National-racial group of subject	Per cent of choices								
	Cauca- sian	vs.	Orien- tal	Part- Haw'n	vs.	Orien- tal	Part- Haw'n	vs.	Cauca- sian
<i>Boys' choices of boys</i>									
Part-Hawaiian	50%		50%	28%		72%*	30%		70%*
Other Minorities	36		64*	50		50	78		22
<i>Girls' choices of girls</i>									
Part-Hawaiian	41		59*	42		58*	54		46
Other Minorities	45		65*	32		68*	48		52
<i>Boys' choices of girls</i>									
Part-Hawaiian	44		56*	46		54	59		41
Other Minorities	38		62	55		45	60		40
<i>Girls' choices of boys</i>									
Part-Hawaiian	49		51	55		45	52		48
Other Minorities	49		51	36		64	45		55

*Per cent. subject

*Per cents which are significantly different from 50 per cent, or chance.

have been classified in two groups. The Part-Hawaiians comprised 13 per cent of the class enrollment and included children who had any Hawaiian ancestry. The remainder of the children (4.3 per cent of the enrollment) have been included in a group designated as "other minorities." Since the majority of the children in both groups are of mixed national and racial ancestry, their preferences are of interest. The small number of children involved, however, means that conclusions from the data must be tentative.

When the Part-Hawaiians were asked to choose between Caucasian and Oriental classmates, neither race was favored consistently. The boys of the two races were chosen equally often, but the Oriental girls were favored over the Caucasian girls.

When the choices of the Part-Hawaiians were from Part-Hawaiian-Oriental and Part-Hawaiian-Caucasian pairs, they did not show any bias toward their own group. However, a real preference for Oriental and Caucasian classmates over Part-Hawaiians was evident only when boys chose boys and for Orientals over Part-Hawaiians when girls chose girls. Whether this failure of Part-Hawaiians to show a favorable bias toward classmates of similar national-racial backgrounds is attributable to their minority status in these particular classes or to their mixed racial backgrounds can not be determined from these data. It may be that both factors were operative.

8. *Attitude of Major Groups Toward Part-Hawaiians and Other Minorities*

The extent to which the Caucasians, Chinese, and Japanese selected the Part-Hawaiian children and children of the "other minorities" group when they were paired with Caucasian and Oriental children is shown in Table 8.

The Chinese and Japanese boys and girls tended to favor Part-Hawaiians over the Caucasians and to favor the Orientals (children of their own race) over the Part-Hawaiians. The Caucasian boys preferred the Part-Hawaiian boys over the Orientals and the Part-Hawaiian girls over both the Caucasian

TABLE 8
STATUS OF PART-HAWAIIAN CHILDREN AND CHILDREN OF OTHER MINORITY NATIONALITIES
WHEN JUDGED BY CAUCASIAN, CHINESE, AND JAPANESE CLASSMATES

National- racial group of subject	Per cent of Part- Haw'ns chosen when paired with:		Per cent of other minorities chosen when paired with:	
	Caucasian	Oriental	Caucasian	Oriental
<i>Boys' choices of boys</i>				
Caucasian	53%	63%*	30%*	53%
Chinese	68*	53	59	53
Japanese	61*	33*	59	38*
<i>Girls' choices of girls</i>				
Caucasian	46	51	50	43
Chinese	33*	37*	75*	59
Japanese	61*	38*	62*	34*
<i>Boys' choices of girls</i>				
Caucasian	58*	66*	42	48
Chinese	59*	35*	67	46
Japanese	59*	40*	49	33*
<i>Girls' choices of boys</i>				
Caucasian	52	50	42	39*
Chinese	58	51	61	54
Japanese	48	34*	53	33*

*Per cents which deviate significantly from 50 per cent, or chance.

and Oriental girls. The choices of the Caucasian girls in pairs involving Part-Hawaiians and Caucasians do not differ significantly from chance, which indicates the former have a more favorable standing than did Orientals when paired with the same Caucasians. When the pairs judged by the Caucasian girls consisted of Part-Hawaiians and Orientals, the choices did not differ from chance expectancy.

The position of the Part-Hawaiians seems, then, to lie between the positions of the Orientals and the Caucasians when judged by members of these two races. They are usually preferred to the other race but not over the race of the judge. Perhaps this position is explained by the fact that the Orientals and Caucasians have adopted more of the Hawaiian customs (especially with respect to food, dress, and language) than they have of the customs of each other. Also, the Hawaiians have assimilated many of the Oriental and Caucasian customs by virtue of their mixed racial ancestry.

The attitude of the children of the three major groups toward the few representatives of mixed ancestry and nationalities which are a minor part of the population is difficult to determine with such small numbers of subjects. Only a third of the percentages differ significantly from chance (50%), and the Japanese subjects accounted for more than half of these percentages.

The choices of the Caucasian children when judging pairs of names involving the "other minorities" group and Caucasian and Oriental children did not deviate from chance except in two instances when the latter were favored.

In the response of the Chinese children, the percentages with one exception favored the minority group, and the magnitude of this preference was greater when the other member of the pair was Caucasian than when he was Oriental. However, the only case where the bias is clearly significant is the choice of Chinese girls for girls in the Minority-Caucasian pairs. Although not shown in the table, the Chinese especially favored the Koreans in the miscellaneous group.

The Japanese tended to favor the minority group over the Caucasian, although only one of these percentages differs significantly from chance. When the minority group was matched with Orientals, the preference of the Japanese subjects in every case reliably favored the Orientals.

The Caucasians did not favor the "other minorities" group, although the rejection is not sufficiently large in most instances to justify any generalization about their attitude. The Orientals appeared to favor the "other minorities" over the Caucasians but not over the Orientals.

D. SUMMARY AND CONCLUSIONS

Two hundred and thirty-three fifth-grade children in Honolulu, Hawaii, representing a variety of national and racial backgrounds, were asked to indicate their preferences when each of their classmates of a given sex and of different national-racial ancestry were paired. The responses of the three major groups of subjects (Caucasian, Chinese, and Japanese) were analyzed to determine the magnitude and variability of national or racial self-preference and the effects on the former of the sex of the subject, the race and nationality of the subject, the relative size of the subject's national group in the class, and the sex of the classmates being judged. The attitude of the major groups of subjects toward their classmates of mixed national-racial ancestry and of other minor groups, and the attitude of the latter toward the former were also studied. These analyses provided the basis for the following conclusions.

1. The Caucasian, Chinese, and Japanese children show a preference for classmates of their respective races when the pairs judged represent the Caucasian and Oriental classmates.

2. The tendency is for Oriental children to show a significantly greater racial self-preference than do Caucasian children.

3. Boys and girls do not differ in their degree of national or racial self-preference when choosing classmates of their own respective sexes.

4. In general, the boys and girls show a stronger preference for their own national-racial group when they are making selections among members of their own sex than when they are making selections among members of the opposite sex.

5. Choices of the boys and girls from the same pairs of classmates result in differences in degree of racial self-preference in about half of the comparisons, with the stronger preference for the subject's own race usually occurring when the classmates being judged are the same sex as the judges.

6. The Oriental subjects show a greater self-preference when their nationality is paired with another race (Caucasian) than when their nationality is paired with another nationality of the same race (Chinese or Japanese).

7. No consistent differences in the variability of the magnitude of racial or national self-preference are evident among the various national-racial groups or between the two sexes.

8. Members of a relatively large national group in a class show a higher average national preference and are more likely to prefer classmates of their nationality exclusively when its members are paired with members of another race than when their national group is relatively small.

9. The Part-Hawaiian children do not favor consistently either the Oriental or Caucasian children when they are matched. Unlike the three major groups of subjects, the Part-Hawaiians do not show a favorable bias for other Part-Hawaiians when these children are paired with Caucasians and Orientals. Whether this trend is due to the minority status of these children in their classes or to their mixed racial backgrounds is not clear.

10. Caucasians, Chinese, and Japanese prefer children of their respective races when they are paired with Part-Hawaiians. When Part-Hawaiians are paired with a race to which the judge does not belong, the Part-Hawaiians are favored. In the preferences of Orientals and Caucasians, the Part-Hawaiians tend to hold a position between the two races.

11. Children representing mixed national-racial backgrounds and minority nationalities tend to favor Orientals over Caucasians. These children are neither favored nor rejected over Orientals by Caucasians, whereas Orientals tend to favor them over Caucasians but to reject them in favor of their own race.

REFERENCES

1. CRISWELL, J. H. A sociometric study of race cleavage in the classroom. *Arch. of Psychol.*, 1939, No. 235.
2. KOCH, H. L. A study of some factors conditioning the social distance between the sexes. *J. Soc. Psychol.*, 1944, 20, 79-107.
3. ———. The social distance between certain racial, nationality, and skin-pigmentation groups in selected populations of American school children. *J. Genet. Psychol.*, 1946, 68, 63-95.
4. MORENO, J. L. Who Shall Survive? *Nerv. & Ment. Dis. Monog.*, 1934, No. 58.
5. THURSTONE, L. L. An experimental study of nationality preferences. *J. Gen. Psychol.*, 1928, 1, 405-425.

11002 Ophir Drive (Apt. D)
Los Angeles 24, California

FURTHER OBSERVATIONS ON THE BEHAVIORAL DEVELOPMENT OF NEGRO CHILDREN*

School of Hygiene and Public Health, The Johns Hopkins University

HILDA KNOBLOCH AND BENJAMIN PASAMANICK¹

A. HISTORY OF THE PROBLEM**

Almost all available studies indicate that Negro school children and adults do not do as well on intelligence tests as whites. Apart from the questions of the cultural factors involved in intelligence tests and of what such tests actually measure, and both are still in the field of controversy, there are many studies which indicate that socio-economic variables play a rôle in the depression of scores on standard tests. In partial answer to Darwin's suggestion, which was apparently based on a commonly held belief derived from coarse estimates of gross motor behavior, Klineberg (7) performed an important study. As a test of the theory that environmental and educational handicaps may account for the failure of Negroes to perform as well as whites and for their decline in performance, he has shown that the longer a Negro school age child has resided in the North the better his performance on intelligence testing. Witty (15) has demonstrated that a group of illiterate Negroes in a Special Training Unit in World War II showed the same ability to learn as a group of illiterate whites. A large number of reports are available in the literature noting that when economic class is to some degree equated, the children of white and Negro families in similar classes tend to perform on approximately the same levels. On the other hand, those who believe that the Negro race is innately inferior point to McGraw's study (8) of Negro infants to show that even in the first year, where cultural influences should be at a minimum, the development of Negroes falls behind that of whites (3). They further intimate that members of the lower social classes rightfully belong in those strata because of their endogenous inferiority.

In an attempt to shed some light on the controversy, to elucidate some of

*Received in the Editorial Office on August 2, 1951.

***It would be desirable to test statistically . . . the truth of the often repeated statement that colored children at first learn as quickly as white children, but that they afterwards fall off in progress.* — Charles Darwin, 1881.

¹The authors wish to acknowledge gratefully the help and advice of Dr. Arnold Gesell and the late Dr. Catherine S. Amatruda. The staff of the Visiting Nurse Association of New Haven also gave invaluable assistance.

the problems involved in the study of racial differences, and to distinguish between innate characteristics and the effect of environmental influences, a study of a group of white and Negro infants was undertaken, and the complete results published elsewhere (11). As a preface to this follow-up report, the main methods and findings of the original study may be summarized here briefly.

The study was conducted in New Haven, Connecticut, a typical New England community of 160,000, with a 3.9 per cent Negro population. A total of 53 Negro infants (28 male, 25 female) was examined, approximately 85 per cent of the infants in the age range under consideration who resided in the Nineteenth Ward. The remaining 15 per cent was almost completely accounted for by two prematurely born children who were not included because of the difficulty in estimating true age, and a few children, possibly of a higher economic bracket, who received private pediatric care. The infants were examined at a Department of Health Well-Baby Clinic and came from homes of average or below average economic status, from the group entitled to post-natal care at a nominal fee.

Two examinations were done, at approximately six month intervals, in the period from June, 1944, to July, 1945. The average age on the first examination was 26 weeks (± 12.1), with a range from 16 to 50 weeks; at re-examination, when 34 babies were seen, the average age was 50 weeks (± 13.5), with a range from 40 weeks to 18 months.

Three groups of white infants were examined for comparison: illegitimate infants living in foster homes examined at the Clinic of Child Development in the course of evaluation as candidates for adoption; infants residing in a child-caring institution (this was a carefully selected group who were normal initially or who rose to within the average developmental range after a more or less prolonged period of foster home care prior to adoption); infants from superior homes who were examined at the Clinic of Child Development as candidates for Nursery School. A fourth group inherently available for comparison was the large group of New Haven white children upon whom the Yale Developmental Schedules had been standardized. It was, of course, realized from the outset that it is not possible to equate completely the environmental variables for a group of Negroes with those for any white group, so that the white groups were presented not as rigid controls but rather for purposes of comparison.

The examination methods used in the study were those developed by Dr. Arnold Gesell and his co-workers, designed and clinically applied to appraise all spheres of behavior including intelligence.

The following conclusions were drawn from the study of the four groups of infants:

1. The average New Haven Negro infant of this study is fully equal in behavioral development to the average New Haven white baby.

2. No outstanding characteristic is found which can be called a "racial" difference, with the possible exception of the definite acceleration in gross motor behavior displayed by the Negroes.

3. Female infants, both white and Negro, are slightly accelerated in development when compared to the males.

4. There are no significant associations of development with depth of pigmentation, number of years of parental schooling, and regional origin of the parents of the Negro infants.

5. Environmental impoverishment is shown to have a marked influence upon a group of institutionalized white children.

6. Two environmental factors correlated with development in this study, although playing some rôle, are apparently insufficient, during infancy, to cause deficiencies in development to the degree reported in previous studies.

- a. Children living in poor housing do not do quite as well as those living in better dwellings.

- b. Children with siblings of the preschool ages, on the whole, do not do as well as infants without siblings.

7. A likely explanation for the normal developmental progress displayed by the Negroes of this study may be found in the recently reported important influence of maternal prenatal diet upon offspring. The Negro babies of this investigation were conceived and born during World War II, when the Negro people of New Haven enjoyed a diet which, for the first time, approached that of New Haven whites. This is partially reflected in the heights and weights of the Negro children which exceed those of any previously reported Negro group and approach the white norms.

8. The second examination of the Negro infants shows a significant advance in development. A similar advance is found in a group of white children of superior backgrounds.

9. When development is tabulated for each half-year of life, that of the second half-year is higher than the first. However, that of the third half-year, although higher than the first, is slightly lower than the second half-year. The onset of the depressing influence of exogenous factors upon Negro development might be construed as beginning during this third half-year of life.

B. PROCEDURE

Of the original 53 Negro infants, 44 (20 males and 24 females) were seen for re-evaluation in the six-week period between March 15, 1946, and May 3, 1946. The remaining nine had moved too far out of town to be seen, or had moved without leaving a forwarding address. Three of the 44 had moved to nearby localities, two to New York City and one to a small Connecticut town; all expressed willingness to have the examiner come to visit.

All of the children were examined by one of the authors (HK), who was present at almost all of the previous examinations.

Only 14 of the children were seen at the Well-Baby Clinic, while 30 were seen in their own homes. This was done to insure that a sufficient number of children be seen in the limited time available, and was not due to lack of co-operation on the part of the parents. None of those approached by the Visiting Nurses refused to have their children seen, and the vast majority went out of their way to make arrangements to be present at the examination or to have an adult available so that the children could be seen; many expressed interest in the study as well as in the professional lives of the examiners. The fact that so many of the children were re-examined bears testimony both to the splendid efforts and co-operation of the Visiting Nurse Association and to the genuine interest of the parents in their children.

The three white groups in the original study were not re-examined, both because of time limitations and because of the inaccessibility of the children who were adoption candidates.

Four of the 44 children are not included in the final tabulations, so that the findings are based on the examination of 40 children (19 males and 21 females) with an average age of 24.4 months (± 3.48) and a range from 18½ to 31½ months. One boy, aged 23 months and examined at home, was timid and inhibited, never really adjusted to the examination, had several violent temper outbursts, and hoarded the test materials without exploiting them. One girl, aged 27 months, and seen at home, was extremely inhibited and reserved, making only a few minimal approaches to the test materials; she was completely self-contained, however, with no evidence of disorganization or fragility, and interestingly enough showed the same kind of behavior to a lesser degree on her previous examinations. The mother implied that the white skin of the examiner exerted an influence. A second girl, aged 20½ months and seen at home, was so completely hyperactive and disorganized in her behavior that only a few items of behavior could be elicited opportunisti-

cally, and no quantifiable evaluation of her development made; on previous examinations she showed good drive and attention, even at a time when she was examined at home, with the house in a turmoil. The third girl, aged 20½ months, was seen at the Well-Baby Clinic at the end of a long day. She had coryza and was extremely cross and irritable, refused many situations by saying "I don't want to," and in general was un-co-operative. She showed no evidence of disorganization, however, and although her final performance was clinically fully average, the results were excluded from the tabulations because a precise developmental quotient could not be secured. This nine per cent failure in achieving complete examinations is perhaps somewhat high, but we do not know what the expected success in other similar series of examinations might be. The queries that must be raised are "how large a rôle did the examiner's skin color play during this re-examination; is this the age when perception of racial differences appears?"

The methods of examination again were those designed at the Clinic of Child Development, the details of which are to be found in *Developmental Diagnosis* (5). During a preliminary interview any changes in occupation, marital status, number of siblings, or housing were noted. An inventory of the behavior in the gross motor, fine motor, adaptive, language, personal-social, and emotional fields was made and an interval health history taken. Following the developmental examination, the child was weighed and measured and the pigmentation of the skin noted. Immediately after the conclusion of the entire examination, the behavior was recorded on the Gesell Developmental Schedules, comments were made regarding the adjustment and the quality of the performance, and a clinical appraisal of the developmental level of maturity in each of the five major fields of behavior was made. On the basis of this clinical evaluation, a developmental quotient² was computed for each field of behavior for purposes of comparison with the normal developmental quotient of 100.

The occupations of the fathers were again arranged roughly according to the Barr scale rating of occupational status (Table 1). Fifteen fathers had changed their jobs since the first examination: three had bettered their positions, four retained the same status, and eight were in occupations which were lower than the original rating.

There was only one change in the educational status of the parents; one

²Developmental quotient = $\frac{\text{maturity age}}{\text{chronological age}} \times 100$

TABLE 1
OCCUPATIONS OF FATHERS OF SUBJECTS

Laborer	3
Porter	1
Truck driver	3
Janitor	2
Stoker	1
Rug cleaner	1
Laundry worker	2
Factory worker	10
Clothes dyer	1
Clothes presser	4
Tailor	1
Mail carrier	1
Electrician	1
Machinist	2
Mechanic	1
Shipping clerk	1
Proprietor of tailor shop	1
News agent	1
Pianist	1
Student	1
Unknown	1

father who was a college graduate returned to school for training in another field.

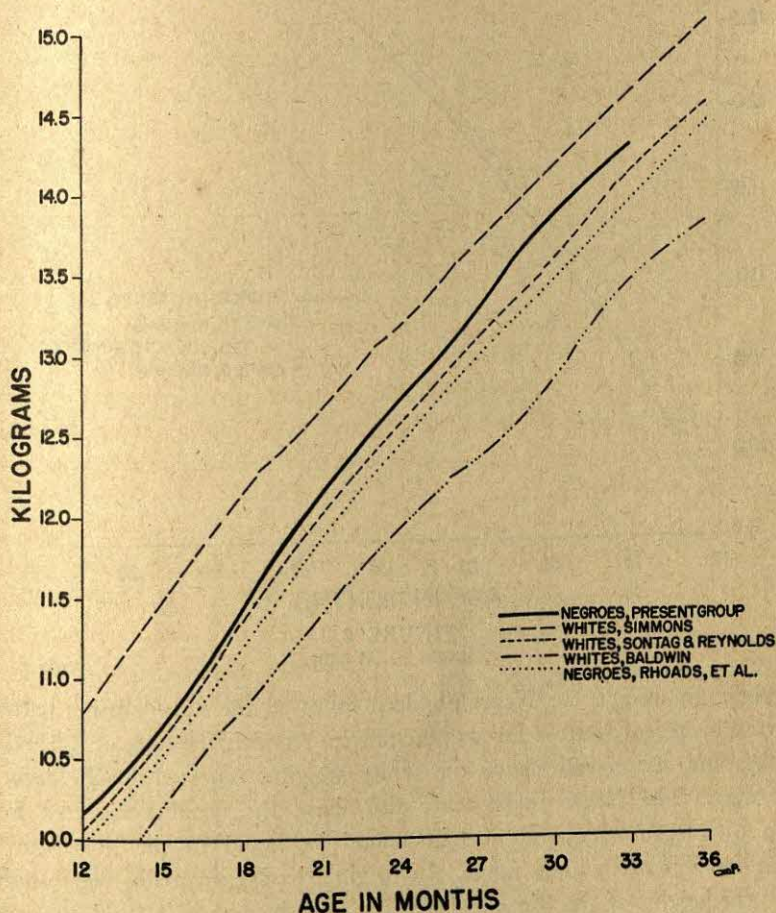
Only six families had moved since the first study. Two families had moved from the housing project to typical slum dwellings, three had moved from the project to fair housing facilities, and one had moved from a slum house into the housing project. Thus 12 infants had good housing accommodations, 21 lived in typical slum dwellings and seven lived in housing which was between the other two groups.

The degree of pigmentation of the children was estimated again and subjectively divided into black [8]^a, dark brown [22], light brown [10], and very light brown [0]. Ten children had not been classified previously since they were not seen at the time of the second examination; three previously called very light brown were reclassified as light brown; seven changed from light to dark brown, and seven from dark brown to black. Thirteen children showed no change in the classification of intensity of pigmentation.

Eleven new siblings were born in the interval between the two studies. Seven were born where the children had previously had siblings; one new sibling was one year old, one five months old and two one month old. Thus 23 children had siblings and 17 had none.

^aAll numbers in brackets indicate number of subjects.

Each child was weighed and measured with indoor clothing but without shoes and stockings; a bathroom scale and a tape measure, with the child standing against a wall, were used for those children examined at home. In addition, the well-baby clinic record was reviewed and the heights and weights at various ages recorded. These weights and heights were averaged by sex for each age interval, plotted on a graph and a smoothed curve drawn. The growth progress was then compared with white and Negro norms as reported in the literature (6a-k). In Figures 1-2 only the highest and lowest curves reported in the recent literature for white children have been plotted for clarity in presentation. The present group fell between these two



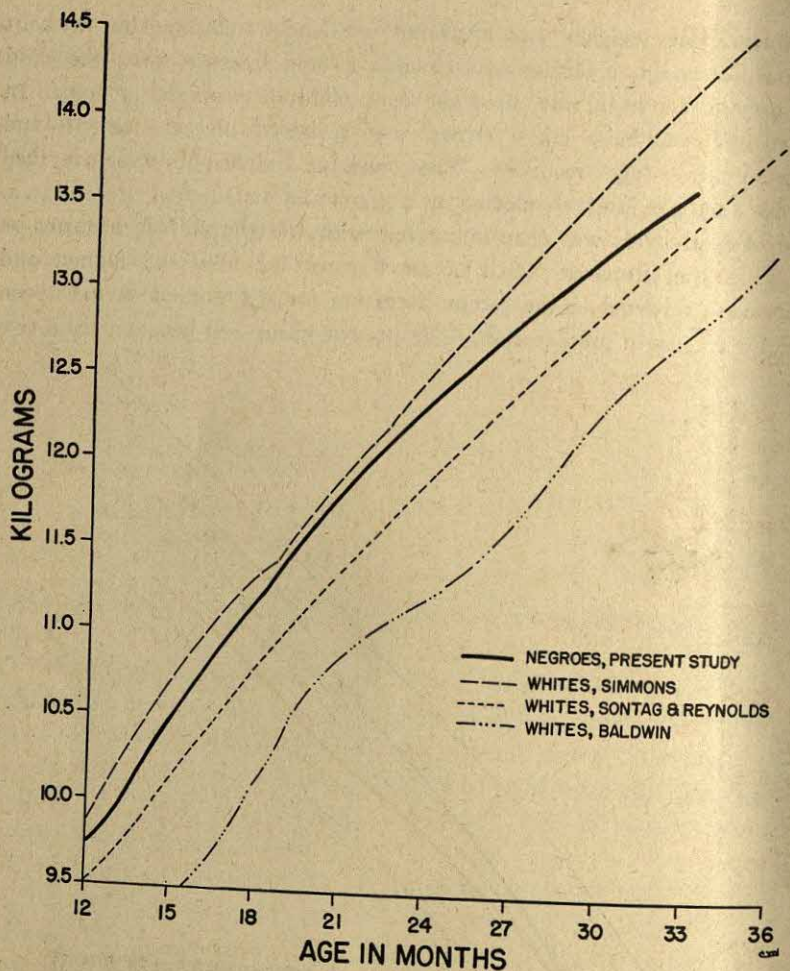


FIGURE 2
FEMALES — WEIGHT

extremes. In weight, the Negro children followed the same growth patterns reported by recent sources for white children between the ages of 12 and 36 months, and were well above the white weights reported by Baldwin at those ages. The Negro males were also above the weights reported for a group of Negro males; no weights could be found in the literature for Negro females of this age range. In height, the Negro males were slightly below the heights of the white children between the ages of 12 and 36 months reported in recent literature, but above those reported by Baldwin and

Woodbury. They made the same progress in height as a reported group of Negro children did. The Negro females showed height curves similar to those for the recent groups of white children but were well above those reported by Baldwin. They were well above the groups of Negro children reported in the literature. Thus, this group of Negro children maintained the growth patterns they showed in the first year of life.

The Well-Baby Clinic record was reviewed and the mother interrogated for a history of illnesses which occurred since the previous examinations. There was a relatively minor amount encountered. Eight children had had

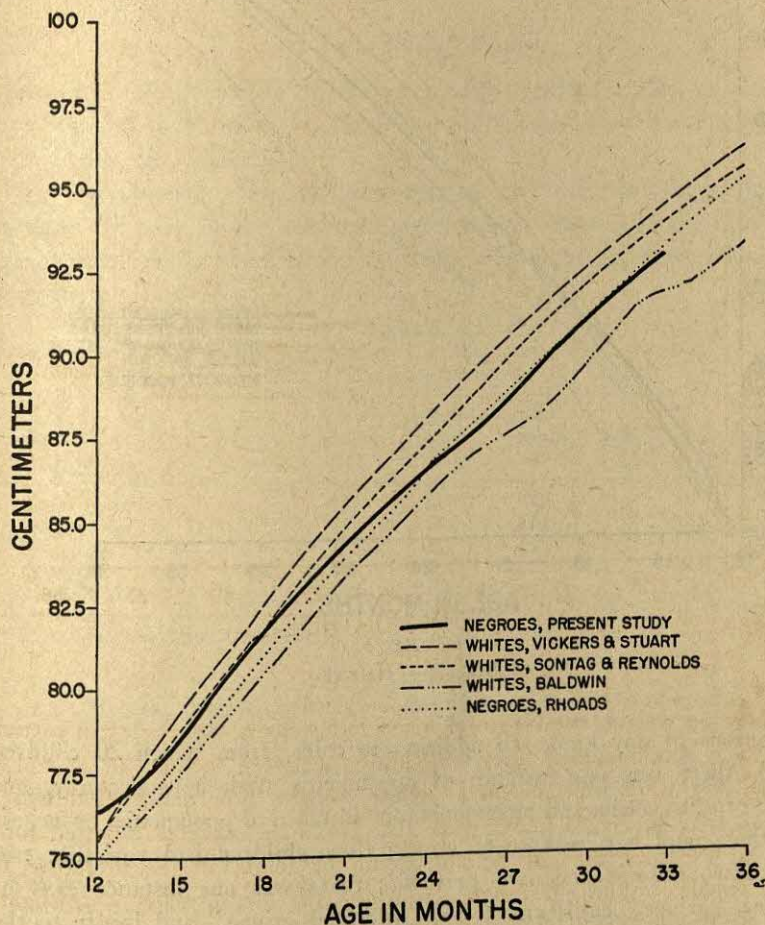


FIGURE 3
MALES — HEIGHT

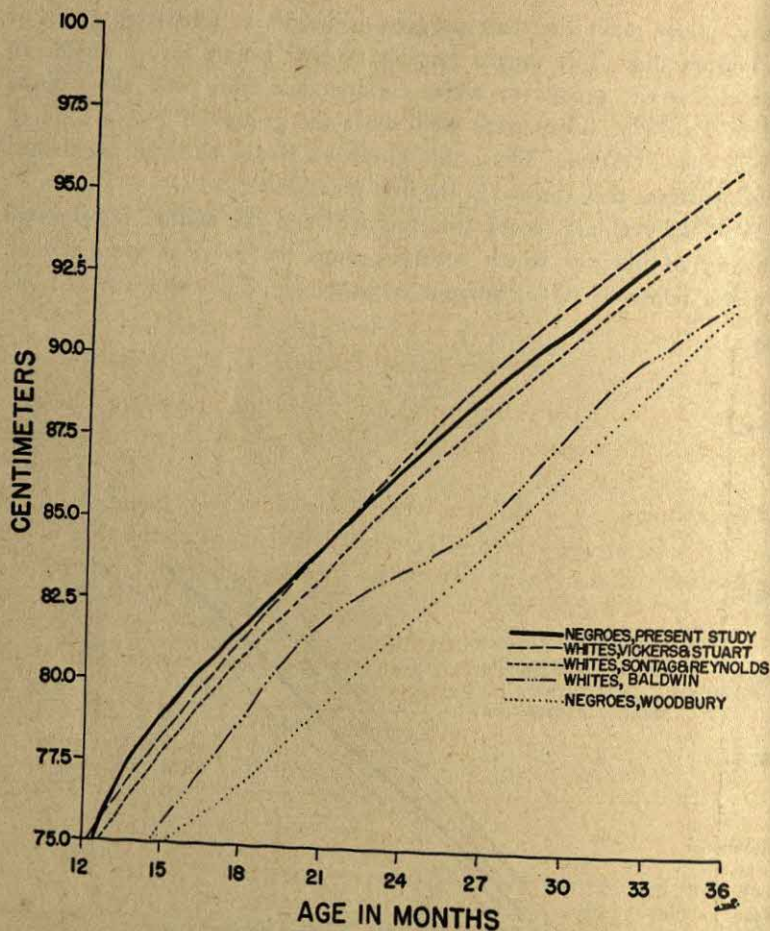


FIGURE 4
FEMALES — HEIGHT

no infections of any kind. In addition to colds, from which 20 children suffered, there was one instance of appendicitis with appendectomy, one instance of otitis media and pneumonia, one instance of pneumonia, bronchitis and asthma, and one instance of croup; three children had varicella, two measles, two mumps and two pertussis; there was one instance each of diarrhoea, "grippe," chills and fever, "intestinal grippe," and injury to the eyelid with infection.

C. RESULTS

1. *Total Group Compared With the Norms*

When the results for the entire group are tabulated for each of the fields of behavior (Table 2), it is seen that this group of Negro children is fully average, with significant and definite acceleration in gross motor development, while language lags significantly behind the other fields of behavior at below the 1 per cent level. [Student's "*t*" test: $P = .01$ In other words, the probability that differences as great as this or greater would be due to chance is less than one in one hundred (2)].

2. *Intra-Racial Findings*

a. Sex. While female infants are usually reported to develop somewhat more rapidly than males, no significant differences are found between the sexes at this age (Table 3).

b. Pigmentation. The children were divided into two groups, light and dark, on the basis of the degree of pigmentation assigned at the time of the examination. Again there is no statistically significant difference between the two groups (Table 4 *a*).

TABLE 2
DEVELOPMENTAL QUOTIENTS OF NEGRO CHILDREN ON FIRST, SECOND, AND THIRD EXAMINATIONS

Field of behavior	First exam. (53)	Second exam. (34)	Third exam. (40)	Standard Deviations, third examination
Gross motor	109	114	124	18.7
Fine motor	102	106	108	11.1
Adaptive	102	107	106	7.7
Language	106	104*	101	10.8
Personal-social	105	108	114	13.2

*Previously reported erroneously as 107

TABLE 3
DEVELOPMENTAL QUOTIENTS OF NEGRO MALE AND FEMALE CHILDREN

Field of behavior	First exam. (28)	Males Second exam. (18)	Third exam. (19)	First exam. (25)	Females Second exam. (16)	Third exam. (21)
Gross motor	108	112	123	110	116	126
Fine motor	102	104	110	102	107	107
Adaptive	101	104	107	103	110	106
Language	106	104	101	105	103	101
Personal-social	106	108	112	103	108	117

TABLE 4
DEVELOPMENTAL QUOTIENTS OF NEGRO CHILDREN ACCORDING TO:

	a) Skin Color		b) Education of parents		c) Geographic origin of parents		d) Birth weight		e) Presence of siblings		f) Housing	
	Light	Dark	Above median schooling	Below median schooling	North	South	Below median birth weight	Above median birth weight	With siblings	Without siblings	Good	Poor
<i>Number of children</i>	10	30	19	21	13	9	18	22	23	17	19	21
<i>Field of behavior</i>												
Gross Motor	125	124	122	126	125	132	115	131	125	122	130	119
Fine Motor	112	107	108	109	108	111	110	107	106	112	108	108
Adaptive	109	105	106	107	107	108	104	109	106	107	109	105
Language	100	101	101	101	100	103	100	101	100	101	102	99
Personal-social	115	114	112	117	112	118	110	118	115	113	116	113

c. Education of parents. Using the median of 10.5 years of schooling, the children were divided into two groups on the basis of whether the parents received more or less than the median schooling. There is little difference between the two groups (Table 4 *b*).

d. Geographical origin of parents. The developmental quotients, when correlated with the geographical origin of the parents show no significant differences between those of Northern and Southern origin (Table 4 *c*).

e. Birth weight. The children were divided by the median into two groups on the basis of birth weight, the heavier group weighing on the average 779 grams more than the lighter. Those children heavier at birth are advanced in all fields except fine motor behavior. In adaptive behavior the difference between the two groups is significant between the 2 per cent and 5 per cent levels (Table 4 *d*).

f. Present weight. When the children were divided into two groups again on the basis of whether, at the time of the third examination, they were above or below the average age-specific and sex-specific weight for the experimental group, as determined from Figures 1A and 1B, those children who were above the mean weight are significantly advanced in adaptive behavior at below the 1 per cent level (Table 5). However, it was observed

TABLE 5
DEVELOPMENTAL QUOTIENTS IN ADAPTIVE BEHAVIOR OF NEGRO CHILDREN ACCORDING TO
THEIR WEIGHT AT THE THIRD EXAMINATION AND AT BIRTH

	Number of cases	Adaptive behavior
Below average weight* at 3rd exam.	20	103
Above average weight at 3rd exam.	20	110
Below median weight at birth and below average weight at 3rd exam.	11	101
Above median weight at birth and above average weight at 3rd exam.	13	111
Below median weight at birth and above average weight at 3rd exam.	7	107
Above median weight at birth and below average weight at 3rd exam.	9	105

*Average weight = average age-specific and sex-specific weight.

that not all children heavier at birth remained above the average weight⁴ at the time of the third examination. Therefore the entire group was divided into four sub-groups. Those children who were below the median weight at

⁴Average weight = average age-specific and sex-specific weight for the experimental group

birth and below the average weight at the time of the examination have the lowest single score in adaptive behavior, while those who were above weight at both times have the highest adaptive behavior quotient in any grouping. The difference between the developmental quotients for these two groups is statistically significant below the 1 per cent level. The difference between the adaptive behavior quotients for those children who were above the median weight at birth as well as above the average weight at the third examination and those who were above the median at birth but whose nutritional status had deteriorated enough to place them below the average weight at the time of the third examination is significant between the 1 per cent and 2 per cent levels. The difference between the children who were below weight at both times and those who were below the median at birth but improved in their nutritional status sufficiently to be above the average weight at the time of the third examination is not significant, lying between the 10 per cent and 20 per cent levels (Table 5).

g. Occupation of fathers. Because a large number of parents were classified in the intermediate group of factory or defense workers with no precise description of their work obtainable, they were disregarded and the remaining 31 were classified as either definitely above or below the median in occupation. The adaptive behavior developmental quotient for those in the below the median occupations is 105, for those above it is 109. No information on actual earnings is available. While the *P* value for the difference between the two groups is only between 0.1 and 0.2, it would be interesting to investigate a larger sample to see if this trend is significant.

When occupation is correlated with housing, 6 of the 16 with below median occupations had good housing and 10 poor housing; 9 of those with above median occupations had good housing and 6 poor.

h. Compared with self (progress.) When the results for the entire group for the first, second, and third examinations are compared, it is evident that the acceleration of development of the second examination over the first is maintained, except in the field of language; in gross motor behavior this acceleration is particularly marked and of even greater degree than was previously present (Table 2).

i. Siblings. When those children who have no siblings are compared with those who have, there is no significant difference between the two groups, although those children without siblings show acceleration in fine motor behavior (Table 4 *e*).

j. Housing. When the children living in good houses are compared with

the children living in poor housing, those with good surroundings are accelerated in practically all fields of behavior; the difference in the adaptive behavior quotients is not statistically significant, lying between the 5 per cent and 10 per cent levels (Table 4 f).

3. *Notes on Personality of Negro Children*

A brief comment on the less tangible factors of personality development is indicated. With the exception of the instances already mentioned, all of the examinations, once begun, were carried out to a successful conclusion. Crying and fussing were negligible, and in the main the group was even more friendly than it had been on the two previous examinations. Behavior in general was stable, mature, and well-controlled, with only a few instances of either exuberance or marked inhibition. It was mentioned previously that one mother inferred that her child's rather unusual inhibition might have been due to the examiner's different skin color. It is still impossible to state with any certainty how important a variable this is at this age level, but the impression was gained that it played a definite although probably slight rôle, but certainly exerted more influence than it had the year previously. This impression has some confirmation in the developmental quotient of 101 for language behavior which is significantly lower than that of the other spheres of behavior at below the 1 per cent level, and also below that found on the first two examinations. Clinically, however, whenever there was a disparity between the various phases of language development, comprehension and reported behavior were almost universally more advanced than behavior requiring a verbal response from the child. The most important question this latter observation raises is the consideration of how significant a factor speech should be in the evaluation of reported scores of the testing of Negroes where the tests have a large verbal component. There were no detectable basically "racial" characteristics which set this group of children apart from any similar group of white children.

D. DISCUSSION

The general conclusion which can be drawn from the data which have been presented is that this group of Negro children has maintained a normal rate of development and continues to be somewhat more advanced than the group of New Haven white children from whom the Yale Developmental Schedules are derived. The common "stereotypes" of ineptness in the handling of fine materials and decline in developmental rate with the failure of "intelligence" to keep pace with motor function are not supported by the present findings.

While gross motor development is undeniably markedly accelerated, there has been no failure to maintain normal developmental rates in the other fields of behavior. In adaptive behavior, the field which probably approximates most closely that described as "intelligence," the group is fully average and there is a significant number of Negro children who are potentially superior.

When an attempt is made to correlate the intra-racial findings discussed below, which may indicate either manifestations of innate characteristics or the results of some environmental influences, with one exception no highly significant differences are found.

Degree of pigmentation is still considered by some, although erroneously, to be positively correlated with the amount of admixture of white blood, and it is postulated on this basis that the lighter Negro should be superior. No such differences between the two groups of Negro children is found.

Education of the parents used as an index of intelligence, while showing a positive correlation with the intelligence of the children in most studies, is a hazardous procedure when applied to a group of Negroes because of the limited educational opportunities available to the class to which most Negroes belong. In this study, the group of children whose parents had the equivalent of a grammar school education did as well as the group whose parents had the equivalent of a high school education.

Most previous reports show that Northern Negroes perform better on intelligence tests than do Southern Negroes. When development is correlated with the origin of the parents, previous reports are not confirmed, and there is no significant difference between the two groups.

There is no difference between those children who have and those who have not siblings. The previous examinations showed those children without siblings to be somewhat accelerated, possibly because they received more attention from their mothers. This attention apparently may be a necessary stimulus to development in the early months, but it seemingly diminishes in importance with maturity and the ability of siblings to be stimulated by each other.

Satisfactory housing conditions may tend to provide better conditions for development, since there is a trend in favor of those with good housing.

A number of reports have appeared in the literature stressing the correlation between good nutrition at all stages of development and superior physical and mental status (10a-g). It has been previously stated that this group of Negro children was above the usual birth weight norm reported for Negroes, possibly due to the good prenatal care and diet that their mothers received.

This was made possible during World War II years by the improved economic status of the Negroes and the effects of rationing on dietary intake. The beneficial influence of rationing on nutritional status has been reported elsewhere also (1) while Meredith has shown that there is some relationship between height and weight and socio-economic status, at least in older children (9). Growth curves show that the group as a whole has maintained its good nutritional status. The positive correlation between weight and development persists and is most marked in those children who had the best start at birth and who maintained a nutritional advantage with growth. Those children in the low birth-weight group who remained low in weight were significantly retarded in comparison to the general group. The previously offered explanation for the developmental normality of this group of Negro children was the fact that they had a good start due to improved prenatal nutrition. Although weight is not necessarily the best index of nutrition for any one individual, group averages of weight do seem to have validity. The direct relationship of height and weight to nutrition has been accepted for a long time, with an apparent basis in experimental findings as evidenced by the work of Wetzel (14) and Spies (13), among others. The evidence points to the fact that the maintenance of good nutrition is important in aiding optimal development, and the advantage in favor of those who maintained nutrition is statistically significant. Even though all groups are fully normal in development this advantage may represent more than mere hereditary variations in body types. Only further study of this group of children will elucidate this point.

The information on occupation provides a rough but adequate guide to economic status, which has a positive correlation with intelligence in most reported studies. Occupation cannot be used as an index in comparing Negro and white groups, because of the restricted job opportunities for Negroes, and even within the Negro groups study shows that only slight correlation exists between the intelligence of school children and the occupations of their parents (12). Although the difference between the two groups in the present study is not great, a trend in favor of those with the more advantageous economic position is present. When occupation and housing are compared, again a trend appears in favor of those with the better occupational status having better living accommodations. While this would warrant further investigation, it is not possible to state on the basis of statistical tests that occupation and housing are significantly correlated with development.

It is well known that almost all longitudinal studies of Negro child intelli-

gence show a lowering of average score with advancing age. On the basis of the first two examinations of this group, it was found that the average score for the third half-year of life was below that for the second half-year, although above that for the first half-year of life, and it was postulated that the third half-year might be the time when the depressant rôle of environmental factors began to make its influence felt. This hypothesis was not borne out by the third examination, which in addition revealed that an oversight had been made in drawing the conclusion. While adaptive behavior during the fourth and fifth half-years of life is above the mean for the entire group, for the sixth half-year of life (based on only 5 cases) it is considerably below the mean for the group as a whole. When, however, the individual children who fell into this oldest age category were reviewed, it was evident that their average score was below the average score of the group as a whole at all ages, and represented a mere fortuitous grouping of subjects. No deterioration in development has occurred with advancing age, and no depressing environmental factors appear to be operating up to the age of two and one-half years for the group as a whole. Certainly this group of Negro children continues to be fully average in behavioral development and growth patterns, and it would appear biologically unsound to postulate, after a continued normal trend for an average period of twenty-four months, a sudden recession in intellectual function as a "racial" characteristic, as a phenomenon discrete from other aspects of biologic and psychologic growth.

Not all of the possible variables have been studied, and there are undoubtedly additional factors which have not been considered in this group of Negro children. It would appear then that again only further study might inform us of what environmental factors could play the rôle necessary to depress the developmental rate, should any decline occur. It may be that given a good start prenatally, and in the period dealt with in this study, two years postnatally, some of the handicaps of unfortunate environmental conditions in the remainder of the preschool period might be decreased in intensity.

E. SUMMARY AND CONCLUSIONS

Forty-four of an original group of 53 Negro infants were examined for the third time and their developmental progress studied and correlated with various individual and environmental factors.

1. The average New Haven Negro child between the ages of 18 and 31 months remains fully equal in behavioral development to his average white coeval.

2. The definite acceleration in gross motor development in the group of Negro children is increased, but no other characteristic, either developmental or in the sphere of personality, might be interpreted as "racial." Although still well at the average level, the significant decline in language behavior since the first examination might well be due to the examiner's skin color, and may be interpreted as the beginning of procedural difficulties. This may be an important observation in relation to verbal components of intelligence tests.

3. No significant differences exist between the sexes at these ages.

4. There are no significant intra-group correlations with depth of pigmentation, education of parents, or regional origin of the parents.

5. When the environmental factors of siblings, housing, and occupational status of the parents are correlated with development, no retarding effect on development is found, although two trends might be noted.

a. The presence of other siblings in the home does not influence development at this age.

b. The possibility that occupational status of parents and housing are positively correlated with development of children warrants further investigation.

6. The previously offered hypothesis of the influence of good maternal diet on the status of offspring seems to retain its validity. A continuing good post-natal nutrition is reflected by weight and height growth curves approaching the highest white norms. Those children who were heavier at birth and who maintained their above-average nutrition are significantly accelerated over those who were below the median at birth and did not improve their nutritional status.

7. The suggestion in the previous paper that the depressant effect of adverse environmental conditions might start in the third half-year of life was incorrectly made.

8. No evidence for a downward trend in development is found. Further longitudinal study is necessary to determine when adverse environmental factors exert a depressing influence on developmental rate.

REFERENCES

1. ARMATTOE, R. E. G. The pattern of youth, an interim report. *Nature* (London), 1943, 152, 217.
2. FISHER, R. A. Statistical Methods for Research Workers, 5th Ed. Edinburgh: Oliver & Boyd, 1934. Pp. 319.
3. GARRETT, H. E. Negro-white differences in mental ability in the United States. *Sci. Mo.*, 1947, 65, 329-333.

4. GESELL, A. *Studies in Child Development*. New York: Harper, 1948. Pp. x +224.
5. GESELL, A., & AMATRUDA, C. S. *Developmental Diagnosis*. New York: Hoeber, 1941. Pp. xii +447.
6. *Growth Curves*:
 - a. BALDWIN, B. T. Physical Growth of Children from Birth to Maturity. Univ. Iowa Stud. Child Wel., 1921, 1, No. 1. Pp. 411.
 - b. MEREDITH, H. V. Physical Growth from Birth to Two Years: I. Stature. Iowa City: Univ. Iowa Press, 1943. Pp. 225.
 - c. MICHELSON, N. Investigations in the physical development of Negroes: I. Stature. *Am. J. Phys. Anthropol.*, 1943, 1, N.S., 191-213.
 - d. ———. Investigations in the physical development of Negroes: II. Weight. *Am. J. Phys. Anthropol.*, 1943, 1, N.S., 289-300.
 - e. PEATMAN AND HIGGONS. Growth norms from birth to five years. *Am. J. Dis. Child.*, 1938, 55, 1233-1247.
 - f. RHOADS, RAPPOPORT, KENNEDY, & STOKES. Studies on the growth and development of male children receiving evaporated milk: I. *J. Ped.*, 1941, 19, 169-189.
 - g. ———. Studies on the growth and development of male children receiving evaporated milk: II. *J. Ped.*, 1945, 26, 415-454.
 - h. SIMMONS, K. The Brush Foundation Study of child growth and development: II. Physical growth and development. *Monog. Soc. Res. Ch. Devel.*, 1944, 9, No. 1, Pp. 87.
 - i. SONTAG, L., & REYNOLDS, E. The Fels Composite Sheet: I. A practical method for analyzing growth progress. *J. Ped.*, 1945, 26, 327-335.
 - j. VICKERS, V. S., & STUART, H. C. Anthropometry in the pediatrician's office: Norms for selected body measurements based on studies of children of North European Stock. *J. Ped.*, 1943, 22, 155-170.
 - k. WOODBURY, R. M. Statures and Weights of Children under Six Years of Age. U. S. Dept. of Labor, Children's Bur. Pub., No. 87. 1921. Pp. 117.
7. KLINEBERG, O. *Race Differences*. New York: Harper, 1935. Pp. ix +367.
8. MCGRAW, M. B. A comparative study of a group of Southern white and Negro infants. *Genet. Psych. Mono.*, 1931, 10, 1-105.
9. MEREDITH, H. V. Stature and weight of children of the United States with reference to the influence of racial, regional, socio-economic, and secular factors. *Am. J. Dis. Ch.*, 1941, 62, 909-932.
10. *Nutrition and Development*:
 - a. ANTONOV, A. N. Children born during Leningrad siege. *J. Ped.*, 1947, 30, 250-259.
 - b. BURKE, B. S., BEAL, V. A., KIRKWOOD, S. B., & STUART, H. C. Nutrition studies during pregnancy. *Am. J. Obs. & Gyn.*, 1943, 46, 38-52.
 - c. KUGELMASS, I. N., POUILL, L. E., & SAMUEL, E. L. Nutritional improvement in child mentality. *N. Y. St. J. Med.*, 1944, 44, 2604-2605.
 - d. People's League of Health, Interim Report. Nutrition of expectant and nursing mothers. *Lancet*, 1942, 2, 10-12.
 - e. SANDELS, S. The use of the Terman-Merrill Scale for measuring intelligence of preschool children. *Tidskr. Psykal. Pedagog.*, 1942, 1, 65-72.
 - f. SMITH, C. A. Effects of maternal undernutrition upon the newborn infant in Holland. *J. Ped.*, 1947, 30, 229-243.
 - g. TOMPKINS, W. The clinical significance of nutritional deficiencies in pregnancy. *Bull. N. Y. Acad. Med.*, 1948, 24, 376-388.
11. PASAMANICK, B. A comparative study of the behavioral development of Negro infants. *J. Genet. Psych.*, 1946, 69, 3-44.

12. ROBINSON, M. L., & MEENES, M. The relationship between test intelligence of third grade Negro children and the occupations of their parents. *J. Negro Educ.*, 1947, **16**, 136-141.
13. SPIES, T. D., & DREIZEN, S. The effect of milk supplements on the growth of children with nutritive failure. *J. Ped.*, 1949, **34**, 393-413.
14. WETZEL, N. C. Assessing the physical condition of children. *J. Ped.*, 1943, **22**, 82-110; 208-225; 329-361.
15. WITTY, P. New evidence on the learning ability of the Negro. *J. Abn. & Soc. Psych.*, 1945, **40**, 401-404.

School of Hygiene and Public Health
The Johns Hopkins University
615 North Wolfe Street
Baltimore 5, Maryland

THE EFFECT OF CLASSROOM ORGANIZATION AND GUIDANCE PRACTICE UPON THE PERSONALITY ADJUSTMENT AND ACADEMIC GROWTH OF STUDENTS*

Department of Testing and Instructional Research, Dearborn Public Schools

JOSEPH JACKSON

A. THE PROBLEM

Certain types of classroom organizational plans such as the homeroom, the departmental, or the block-of-time permit a greater or lesser concentration of time for the teacher about her pupils. As a result, opportunity for the study of the intimate problems of children on behalf of a teacher is increased or decreased, depending in part upon the type of classroom administrative philosophy practiced within a building or even within an academic level such as elementary, junior high, or senior high. May one infer, therefore, that if pupil contact about one teacher is increased or decreased, student maladjustment will be directly or inversely commensurate with such contact? If such were to be the case, the homeroom organizational pattern would occupy an eminent position in the guidance philosophy of the classroom, or conversely speaking, the departmental classroom organizational pattern would be the least desirable plan for approaching the emotional upheavals of children.

To study this problem, two separate experiments were devised and carried on concurrently with somewhat different approaches, both yielding a common comparison. For purposes of clarity, I shall present the analyses separately; first, a summary of the Lowrey Junior High School experiment and second, a summary of the Maples Elementary and Junior High School experiment.

B. DEFINITIONS OF TERMINOLOGY

1. *The Homeroom Plan*

Under the homeroom organizational plan, a majority of the pupils' subjects were taught by the same teacher. In this study, the homeroom activities included a daily 15-minute record period and the teaching of English, mathematics, and social studies. Health education, art, and music were taught by separate teachers. The homeroom organizational plan provided a maximum pupil contact for one teacher and the guidance initiative rested with her.

*Received in the Editorial Office on August 7, 1951.

The homeroom organizational plan was considered as one experimental variable in this study.

2. *The Block-of-Time Plan*

In this plan, pupil contact about the same teacher is characterized through a maximum two-hour or two-subject classroom association plus a 15-minute daily record period. The remaining subjects were taught by separate teachers. The two-hour block-of-time organizational plan is a compromise between the homeroom and departmental plans. Guidance initiative rested with the teacher in charge of the block-of-time classroom plan. Mr. *A* was in charge of this plan the first semester, and Mrs. *B* for the majority of the remaining two semesters. The block-of-time plan was considered as a second experimental variable in this study.

3. *The Departmental Plan*

Under this plan the pupil encountered a different teacher hourly for each subject. From the standpoint of guidance philosophy, the departmental plan is the least desirable, since responsibility for classroom counseling and follow-up is dependent totally upon the initiative and good will of some indeterminate person. Mrs. *C*'s contact with her group was limited to one class daily, plus the daily morning 15-minute record period for the first 20 weeks of the experiment and to only the 15-minute record period for the remaining 40 weeks. The departmental approach was considered as the control variable of the study.

4. *Some Additional Variants of the Study*

During the one and a half years of experimentation, some variants introduced into the Lowrey Junior High School study turned out to be highly desirable. Originally, it was planned to retain the same teachers for the duration of the experiment at the Lowrey School; however, the block-of-time group was actually under the jurisdiction of three teachers, shifts having been made because of scheduling and of illness. In the case of the homeroom group, it was also planned to create an unusually strong guidance situation about the pupils, providing them with an opportunity for various psychological and counseling services, pupil planning, and the like. Such plans, however, did not materialize because of the homeroom teacher's assignment to a new camping program, her multiple subject preparations, core curriculum committee work, as well as her graduate load at a nearby university. No doubt variants of the nature described creep into most classrooms at one time or another; therefore, they help to portray the true character of a teacher's

work load. Is it possible that such conditions may affect pupil-teacher relationship from the standpoint of adjustment? Certainly they should be considered in a study of this nature.

C. THE PROCEDURE

1. Controls

If student adjustment is the result of the opportunity for guidance or nonguidance growing out of teacher-pupil contact, certain variables had to be controlled in this experimental situation. Controls were exerted as follows:

1. Equation for the intra-class situations was possible on the basis of gender, intelligence, mental age, teacher, and subject matter time allotment. At the Lowrey Junior High School, the homeroom, block-of-time, and departmental groups were equated while in Grade 7A. At the culmination of the experiment a year and a half later, these groups were in Grade 8A. Of the original 23 students paired in each group, 15 remained for the entire three semesters of study.
2. A reasonable equation of student adjustment was obtained from personnel cards, sociograms, personality tests, and teacher ratings.
3. A test-retest situation was utilized in the areas of personality adjustment, sociometry, and academic achievement.
4. The course content and the teachers involved were agreed upon. The experimental groups, namely, the homeroom and the block-of-time, were to introduce the following variations into the study:
 - a. Additional time and techniques necessary for good guidance practices, such as case studies, group therapy, personal interviews, follow-ups in the home, etc.
 - b. Access to psychological and counseling services within the school system.
 - c. Emphasis upon the study of emotional problems through intimate understanding of needs and of abilities through sociometric and personality testing and guidance.
5. The Null Hypothesis and the Standard Error of the Difference were employed as tests of significance in this study.

2. *An Analysis of Sociometric Variance*

The analysis of variance for the departmental group, Mrs. C's group, yielded the only positive significant result among the three organizational patterns evaluated at the Lowrey Junior High School. The social compatibility

ty had doubled for the departmental group during the sociometric test-retest interim, while the individual status indices increased 229 points. See Tables 1 and 2 for the comparisons. The analyses of these results yielded a significance value beyond the 1 per cent level for differences between tests, as well as between techniques employed. The departmental group was characterized in organization by a daily 15-minute record period and a one semester course

TABLE 1
FLUCTUATION IN SOCIOMETRIC RESULTS FROM TEST TO RETEST, LOWREY SCHOOL

GROUP I—Block-of-time		
Experimental		
Compatibility Index	Tested 2-49	Tested 1-50
Status Index Range	.13	.09
Change in Status Index Negative	—37 to +33	—31 to +34
Change in Status Index Positive	14 decreased a total of 151 points	
Net Change in Status Indices	12 increased a total of 87 points	—64
GROUP II—Departmental		
Control		
Compatibility Index	Tested 2-49	Tested 1-50
Status Index Range	.06	.12
Change in Status Index Negative	—46 to +29	—14 to +41
Change in Status Index Positive	7 decreased a total of 54 points	
Net Change in Status Indices	16 increased a total of 283 points	+229
GROUP III—Homeroom		
Experimental		
Compatibility Index	Tested 2-49	Tested 1-50
Status Index Range	.12	.07
Change in Status Index Negative	—53 to +39	—53 to +29
Change in Status Index Positive	14 decreased a total of 188 points	
Net Change in Status Indices	9 increased a total of 99 points	—89

TABLE 2
ANALYSIS OF SOCIOMETRIC VARIANCE AT THE LOWREY SCHOOL

	DF	Sum of squares	Mean square	F	Expected 1%	Expected 5%	Hypotheses	Difference
GROUP I—Block-of-time								
Experimental								
Test-Retest	25	.7984	.0319	4.62	2.66	1.98	rejected	significant
Between Techniques	1	.0079	.0079	1.25	7.77	4.24	accepted	insignificant
GROUP II—Departmental								
Control								
Test-Retest	24	1.1641	.0485	4.53	2.66	1.98	rejected	significant
Between Techniques	1	.1049	.1049	9.80	7.82	4.26	rejected	significant
GROUP III—Homeroom								
Experimental								
Test-Retest	23	1.1586	.0504	4.09	2.66	1.98	rejected	significant
Between Techniques	1	.0165	.0165	1.35	7.88	4.28	accepted	insignificant

in general science, both allotments being the maximum contact with the sponsor of the departmental group. Despite this time limitation, teacher influence upon social adjustment, as shown through a sociometric evaluation, transcended the influence of the organizational pattern of the teaching situation. This is readily understood through the personality of the sponsor of the departmental group, who was chosen for the experiment for this reason.

The variance for the block-of-time group and for the homeroom group had shown a drop in sociometric compatibility from test to retest, despite an increase in pupil-teacher contact. The block-of-time pupil-teacher contact was characterized by a two-hour or a two-subject block-of-time plus the usual 15-minute daily record period. The homeroom group received ample opportunity for pupil-teacher contact in part through the teaching of social studies, English, and mathematics. In addition to this time, pupil-teacher contact was maintained through a 15-minute daily homeroom period. The block-of-time group had nearly 250 per cent more opportunity for guidance contact with its sponsor than did the departmental group, while the homeroom group had about 350 per cent more opportunity for guidance contact with its sponsor than did the departmental group. The above data are portrayed in Tables 1 and 2.

The sociometric variance for the block-of-time and for the homeroom group indicated similar trends. The retest after a year's lapse had shown a decrease of $-.64$ and $-.89$ points respectively in pupil status indices. This negative trend indicated less social integration rather than more. Is this negative change significant? The F tests of 4.62 and 4.09 for the negative trends from test-retest of block-of-time and homeroom groups show that drops in the compatibility indices were due to more than chance factors.

3. *Differences in Achievement on the California Test of Personality*

While differences between group achievement on the California Test of Personality fail to disclose any critical ratio as large as 3.0, nevertheless, some pertinent trends are noted in the final results. The departmental group under Mrs. C, which had shown a significant difference at the 1 per cent level in the sociometric evaluation, shows through the group paper-pencil test summary a much better adjustment score than the homeroom group. The critical ratio between these two comparisons is 2.69, which indicates that the difference is significant at the .72 per cent level of confidence. The comparison between the group paper-pencil results of the block-of-time and the departmental group indicated almost the same difference; the critical ratio in the latter case was 2.42.

The result of group paper-pencil personality test and of the sociometric survey because of their consistency lead to a number of interesting observations about the environmental setting within each organizational plan. Mrs. C, the teacher of the departmental group, has always generated a warm personal feeling for the needs of her pupils. Three semesters of sponsorship of the departmental group in this study had provided her with maximum opportunity to meet the social and personal wants of the students. The two-hour or two-subject block-of-time plan failed to materialize beyond the first semester into a strong guidance situation because of shifting of teachers in charge of the block-of-time plan, two of whom were unprepared to participate in the study to the degree expected. The results, perhaps, mirror the instability surrounding such a teaching situation. The homeroom plan as used at the Lowrey School was unable to function as a strong classroom situation for a number of reasons. The teacher in charge had multiple preparations in subject matter not only during one semester, but for each of the two succeeding semesters. The sponsor of the homeroom group was also a member of the camping and core curriculum committees and was attending graduate classes at a nearby university. One might rightfully surmise from the sociometric and personality data analyzed that extraneous impositions upon the students, such as teacher turn-over, teacher work load, and excessive committee work can terminate with an impersonal teacher-pupil relationship. A classroom organizational pattern such as the homeroom or the block-of-time type will hardly compensate for such internal interferences, interruptions, and distractions. Administratively, it becomes necessary to evaluate teacher personalities in light of their work loads before assigning them to such a personal, time-consuming undertaking as pupil guidance.

D. NATURE OF ACADEMIC DIFFERENCES ON PRE-TESTS AND FINAL TESTS

1. *Pre-Test Ability Differences*

The mean intelligence quotient for the homeroom and the block-of-time groups at the Lowrey School is 112 and for the departmental group it is 111. As far as learning aptitude is concerned, unusually good pairings existed among the three study groups.

The pre-test differences between mean achievement on the Iowa Work-Study Skills, Iowa Silent Reading, the Stanford Arithmetic, the Stanford General Science, and the Stanford Social Studies, revealed one value with a critical ratio of 3.0. The critical ratio of 3.0 was found in arithmetic achievement between the homeroom and the departmental groups, which

portrayed the departmental group as being significantly superior in arithmetic ability. The remaining critical ratios are considerably under the significance value of 3.0 and thereby reflect chance variations among the test results for group differences on work study skills, reading, general science, and social studies.

2. *Final Ability Differences*

No critical differences of 3.0 or more were found in work study skills, silent reading, arithmetic, general science, and social studies achievements at the conclusion of the experiment three semesters later. Therefore, one might infer that whatever differences in academic abilities remained between groups at the conclusion of the experiment were of a chance nature. It appears that time is rather well allotted for academic purposes under each of the organizational patterns; however, the pupil's emotional response and personality development hinges upon other factors, the most important being perhaps the warmth and feeling of the teacher and the stability of the administrative schedule from semester to semester.

E. SUMMARY OF THE MAPLES SCHOOL EXPERIMENT

The homeroom situation as utilized in the Lowrey Junior High School experiment could hardly be called conventional or representative of homerooms throughout the Dearborn Public Schools. The second experiment at the Maples Elementary and Junior High School for three semesters, beginning with Grade 7B, attempted to determine the guidance merits of a strong homeroom situation when compared to that of a conventional departmental classroom plan. In Grade 5A, for the same duration, an attempt was made to determine the guidance effectiveness of a strong homeroom plan as against a more conventional homeroom plan. A strong homeroom plan represented a situation in which emotional and academic needs of the children were of utmost importance, as already described.

The two strong homeroom situations, namely the experimental groups in Grades 7B and 5A, were essentially the same. Teachers adept at homeroom guidance were chosen. The homeroom teachers taught at least three subjects daily to their group and an additional 10 minutes daily was provided for record purposes. The Grade 7B departmental or control counterpart was exposed to a different teacher hourly for each subject during the 60 weeks of the study. Mr. D, who was in charge of the Grade 7B control group, had a contact of 10 minutes daily plus one class for 20 of the 60 weeks of the study. Time and responsibility for group guidance were not fixed through this departmental organizational procedure.

Miss *E*, in charge of the Grade 5A control group, had essentially a conventional homeroom situation as had Mrs. *F*, her experimental partner. The differences between these two homeroom situations were elicited through aggressive guidance and teaching procedures on behalf of Mrs. *F*, in charge of the experimental homeroom group. It should be stated that the teacher influence among all groups was well matched from the viewpoint of interest, initiative, and ability.

F. AN ANALYSIS OF SOCIOMETRIC VARIANCE

Sociometrically, what changes transpired within each group during the three semesters of study? The group compatibility index of Grade 7B dropped .03 from test to retest for the strong homeroom group, and dropped .04 for the departmental control group. The net change, however, of student indices was $-.19$ for the homeroom group and $-.62$ for the departmental group. On the basis of individual student status indices, it appears as if the strong homeroom situation tended for better social integration than did the departmental control group.

Mrs. *D*'s group, the control half of the Grade 7B study, failed to reveal a gain of significant difference from test to retest for the sociometric data. The *F* value of .35 for differences between the techniques evaluated is considerably below the 5 per cent level of significance. It appears that the departmental plan failed to influence a significant change in the sociometric results from test to retest. See Tables 3 and 4 for a more detailed comparison.

What is the outcome sociometrically if two homeroom situations are evaluated, one being conventional in its procedure, the other being experimental within which sundry psychological and teaching techniques were incorporated into the classroom?

The results of the strong homeroom situation in Grade 5A indicate a cohesive sociometric classroom organization at the end of the experiment. A net increase of 72 points was noted in student status indices from test to retest. The *F* test of 2.86 is significant at the 1 per cent level for the test-retest data; therefore, improvements in the student status indices are not due to chance alone. In comparison, the conventional homeroom results have shown a decrease of $-.96$ points for the sociometric data from test to retest. This negative change is conditioned by other than chance factors, since its *F* value of 4.37 is beyond the expected 2.86, or the 1 per cent level of significance. A difference between the techniques employed during this interim is significant. The *F* value of 1.57 is below the 5 per cent value of 4.32 required for significance.

TABLE 3
FLUCTUATION IN SOCIOMETRIC RESULTS FROM TEST TO RETEST, MAPLES SCHOOL
AT BEGINNING AND AT CONCLUSION OF EXPERIMENT

	GRADE 7B Tested 2-49	GRADE 8B Tested 1-50
GROUP A—Homeroom		
Experimental		
Compatibility Index	.14	.11
Status Index Range	—32 to +32	—53 to +42
Change in Status Index Negative	10 decreased a total of 129 points	
Change in Status Index Positive	7 increased a total of 110 points	
Net Change in Status Indices	—19	
GROUP B—Departmental		
Control		
Compatibility Index	.18	.14
Status Index Range	—37 to +48	—16 to +58
Change in Status Index Negative	12 decreased a total of 193 points	
Change in Status Index Positive	6 increased a total of 131 points	
Net Change in Status Indices	—62	
GROUP C—Homeroom		
Experimental		
Compatibility Index	.10	.09
Status Index Range	—54 to +50	—15 to +37
Change in Status Index Negative	13 decreased a total of 203 points	
Change in Status Index Positive	13 increased a total of 275 points	
Net Change in Status Indices	+72	
GROUP D—Homeroom		
Control		
Compatibility Index	.13	.08
Status Index Range	—48 to +48	—50 to +50
Change in Status Index Negative	13 decreased a total of 262 points	
Change in Status Index Positive	7 increased a total of 166 points	
Net Change in Status Indices	—96	

G. DIFFERENCES IN ACHIEVEMENT ON THE CALIFORNIA TEST OF PERSONALITY

The sociometric data provided an analysis of group structure from the pupil's point of view. How do individuals feel about themselves when asked to interpret their own adjustment through a paper-pencil type of personality test? In other words, how do they feel about themselves and their personal needs, which feelings in turn might possibly reflect an opportunity for adjustment under the various classroom organizational plans to which they have been exposed for one and a half years?

The pre-test of group differences as evaluated by the California Test of Personality fails to reveal any significant variations between mean achievement of the experimental and control groups in Grade 7B as well as in Grade 5A.

TABLE 4
ANALYSIS OF SOCIOMETRIC VARIANCE AT THE MAPLES SCHOOL

VARIANCE ANALYSIS AT THE MAPLES SCHOOL								
	DF	Sum of squares	Mean square	F	Expected		Hypotheses	Difference
					1%	5%		
GROUP A—Homeroom								
Experimental								
Test-Retest	18	1.3666	.0759	5.84	3.13	2.22	rejected	significant
Between Techniques	1	.0009	.0009	.07	8.28	4.41	accepted	insignificant
GROUP B—Departmental								
Control								
Test-Retest	19	.9244	.0495	1.79	3.03	2.17	accepted	insignificant
Between Techniques	1	.0096	.0096	.35	8.18	4.38	accepted	insignificant
GROUP C—Homeroom								
Experimental								
Test-Retest	26	1.8285	.0703	2.86	2.56	1.93	rejected	significant
Between Techniques	1	.0110	.0110	4.47	4.22	7.72	rejected	significant
GROUP D—Homeroom								
Control								
Test-Retest	21	2.4729	.1178	4.37	2.86	2.08	rejected	significant
Between Techniques	1	.0419	.0419	1.57	8.02	4.32	accepted	insignificant

What differences existed between the groups at the conclusion of the experiment a year and a half later? On the final test in Grade 7B, the mean score of the experimental group was 78 and that of the control group was 48. The mean difference of 30 divided by its standard error of the difference yielded a critical ratio of 3.01, which incidentally was the only significant difference among the paper-pencil test results. Here again we had some indication of the effectiveness of strong homeroom guidance situations as reflected through responses to a paper-pencil personality test. In Grade 5A, the difference between the final results of the California Test of Personality was one point on the side of the experimental group; however, it should be remembered that the increase between means for the experimental group from test to retest was eight points as compared to four for the control group. A teaching environment sensitive to the personality problems can alter for the better the social and emotional maladjustment of its students. The experimental situations in Grades 7B-8B, and in Grades 5A-6A, point out that classroom consciousness about the personal problems of pupils and activity on behalf of the teacher in individual and group therapy may mitigate the effects of maladjustment within three semesters of pupil-teacher contact or less. Classroom administration is indicted if the organizational pattern fails to prescribe opportunity for such activity.

1. *Pre-Test Ability Differences*

The mean intelligence of the Grade 7B groups is 104 and of the Grade 5A groups is 105, as obtained by the Otis Quick-Scoring Mental Ability Test, with practically identical standard deviations. The Grade 7B-8B academic results indicate slight superiority of the experimental group in the mean achievement on the tests of Iowa Work-Study Skills, Iowa Silent Reading, Stanford Arithmetic, Stanford Elementary Science, and the Stanford Social Studies. The differences in mean academic achievement between the groups, however, are insignificant. This is likewise true of student pre-test achievement differences in the Grade 5A-6A experiment.

2. *Final Ability Differences*

After three semesters of experimentation, a re-test of academic achievement failed to disclose any critical differences between the results of the experimental and the control groups. Whatever academic variations in achievement between groups transpired after three semesters of study can be considered as mainly due to chance factors. Perhaps from these minor differences in academic achievement one might infer that teaching responsibility is well discharged as the primary requisite within a classroom, while the emotional or personal outcomes are largely dependent upon a number of such factors as teacher interest, time, motivation, specialized help, administrative plan, and the like.

H. OBSERVATIONS

1. A summary of classroom practices reveals that the primary responsibility of the teachers is concerned with the provision of academic growth for the students which, on the basis of comparative achievement, seems to be equally well established and independent of any classroom organizational plan.

2. An increase in pupil-teacher contact through an organizational plan such as the homeroom or the block-of-time does not necessarily provide better sociometric or personality integration within a classroom unless the teacher influence provides attention to such pupil needs.

3. Such classroom administrative factors as teacher turnover, multiple preparations, scheduling upheavals, and excessive extra-curricular activities can render classroom guidance ineffective. On the other hand, a departmental plan which provides teacher tenure and a teacher-pupil relationship characterized by warmth and feeling can contribute significantly to the emotional wants of its constituents.

4. Conventional classroom procedure, no matter what its organizational plan, will contribute very little to the emotional wants of the students unless an effort is made to delve into the individual and group dynamics of the situation. As a matter of fact, group sociometric and personality characteristics may continue to deteriorate.

Testing and Instructional Research

Dearborn Public Schools

5757 Neckel Avenue

Dearborn, Michigan

SHORT ARTICLES AND NOTES

The Journal of Genetic Psychology, 1953, **83**, 171-178.

CHILDREN'S CONCEPTS AND STEREOTYPES OF TURK, PORTUGUESE, ROUMANIAN, ARAB, CHINESE, FRENCH- CANADIAN, MULATTO, SOUTH AMERICAN, HAWAIIAN, AND AUSTRALIAN*

Public Schools of Cincinnati, Ohio

ROSE ZELIGS

A. THE PROBLEM

Conditioning the child to the cultural patterns of his group begins with his birth and continues through life. The concepts and attitudes of his associates become his by means of attendant and direct learning. Intergroup prejudice, often the result of stereotype, is communicated to the child through his social contacts. His concepts of different races and nationalities are often based on distorted, false pictures and stereotypes. A knowledge of children's racial or national concepts is helpful in understanding their intergroup attitudes and in giving the educator a starting point in inter-cultural education.

Blake and Dennis (2) asked pupils from Grades 4 through 11 in a Virginia school for white children to compare Negroes with white people in regard to each of 60 characteristics. The younger children showed a generally unfavorable attitude toward the Negro. The high school children showed a high degree of agreement in their stereotypes (1943).

Meltzer (6) presented children from Grades 5 through 8 with a list of 21 races and nationalities and asked them to indicate their feelings by placing a check under intense like, like, neutral, dislike, or intense dislike. The children were then asked to give reasons for their answers. He combined the data obtained in 1934 and 1938 and classified the children's reasons under the most common notions. In his discussion he notes some differences but states that "concepts of the various nations and races given by the children in 1934 parallel those given in 1938."

Negro students studied by Meene (5) had the same racial stereotypes in 1935 and 1942 except that *revengeful*, *cruel*, and *treacherous* were added to

*Received in the Editorial Office on June 15, 1951.

Germans, less favor was expressed toward the Japanese, and more toward the Chinese.

W. S. Gregory (3) studied stereotyped thinking of 50 college students by having them rate 15 nationalities on a 9-point scale. A month later the subjects were asked to identify the nationality of photographs and to rate them on a 9-point scale. The photographs were not identified correctly (1938-39).

Bayton (1), studying the racial stereotypes of Negro college students, found that the Negro and white students "exhibit a high degree of similarity in the stereotype they possess of racial and national groups" (1941).

Kuhlen (4) asked college students to describe their attitudes towards peoples and then judges rated the material according to the number of favorable and unfavorable statements. Kuhlen also listed the words most frequently used by the students to describe various countries (1942).

Shoenfeld (7) studied problems relating to stereotype in relation to racial prejudice of college students. He found "no relation between either positive or negative direction and uniformity. Positive direction tends to be associated with greater familiarity and negative direction with less familiarity" (1942).

The purpose of this paper is to give children's concepts and stereotypes of Turk, Portuguese, Roumanian, Arab, Chinese, French-Canadian, Mulatto, South-American, Hawaiian, and Australian.

The subjects of this study, tested in June, 1931, were 200 sixth-grade, 12-year-old children, with a large percentage of Jewish children, 98 per cent of whom were native-born. More than half their parents were also native-born, and many of them were of German-Jewish and Russian-Jewish descent. There were also Protestant white, native-born children, and a few Negro children. The average chronological age of the children was 12 years and the average mental age, according to the Otis Group Intelligence Test, was 14 years and four months. The socio-economic background, according to the Sims Score Card, was somewhat below "very high." The children attended a Cincinnati suburban public school whose enrollment was approximately 1,600 students in 1931 and somewhat less in 1944. Other subjects were 174 12-year-old children of the same age, grade, school, and background, tested in June, 1944, and another similar 100 subjects of the same school tested in December, 1944. The data for both 1944 groups were combined.

B. PROCEDURE

The children were given the Zelig's Intergroup Attitudes Test. This requires the children to express willingness or unwillingness to have the

relationships of cousin, chum, roommate, playmate, neighbor, classmate, and schoolmate, with each of 39 races and nationalities. To obtain the children's concepts and stereotypes the children were asked to write what they considered true, interesting sentences about each of the races and nationalities listed on the test. Another method was to have the children write a word or phrase describing the appearance, character, or personality of the race or nationality mentioned. The children's concepts for Turk, Portuguese, Roumanian, Arab, Chinese, French-Canadian, and Mulatto were obtained for both 1931 and 1944. Their reactions to South American, Hawaiian, and Australian were obtained only in 1944.

The data were tabulated to get the children's concepts and stereotypes for each race and nationality mentioned. The reactions were grouped under favorable, unfavorable, neutral, and "don't know" concepts.

C. RESULTS

In 1931 there were 14 per cent favorable concepts of Turk. They are interesting people with an interesting history, are strong, fearless, good fighters, are great story tellers, and wear fancy and beautiful clothes.

Forty-six per cent of the reactions were unfavorable. The Turks have dark skin, dress loud, wear funny clothes and towels on their heads, are war-like, dangerous, cruel, mean, and fierce, and they have many wives. They are hard to teach religion and were on the German side in the war.

The 31 per cent neutral concepts said they wear turbans, long, baggy pants, white, long, flowing robes, and dress differently from us. They have different ways of doing things, live in different houses from ours, and have a different religion.

In 1944 six per cent of the concepts were favorable. They said the Turks are good fighters, are O.K., peaceful, well-to-do, civilized, thoughtful, and make beautiful palaces.

Seven per cent unfavorable responses said they are dark-colored foreigners, cruel, war-like, dirty-looking barbarians, who speak a funny language and are not liked. Eighty-five per cent of the associations were neutral. Most of these were, person from Turkey, or turkey. Others were, desert tribe, beards, horse-riders, Asia Minor, Turkish bath, and enemy of Russia.

In 1931 there were 12 per cent favorable concepts of Portuguese. They were great sailors, are good dancers and singers, have plump healthy children, and are nice, well-trained people and good traders. They are fond of Americans and kept out of the war.

The 4 per cent unfavorable remarks are that they have dark skin, have little money, and refused to help Columbus. Sixty-two per cent were neutral concepts. The Portuguese are like the Spanish, and live near Spain. They are sailors, explorers, and have colonies. They raise ducks, cork, olives, make and drink wine, and export much tropical fruit. They have large families. There were 22 per cent "don't knows."

In 1944 eight per cent of the responses were favorable. They said the Portuguese are O.K., neat, nice, look like us, like Americans, and are good fighters. The 2 per cent unfavorable responses described them as brown foreigners with a fat stomach. Fifty-nine per cent of the concepts were neutral. They were persons from Portugal, Portuguese, ports, geese, ducks, Spain, near Spain, Spanish, Columbus, wine, cork, olives and sailors. There were 31 per cent "don't knows."

Ten per cent of the responses to Roumanian were favorable. They love bright colors, are at peace with their neighbors, have many oil-wells, and issue expensive stamps. The 5 per cent unfavorable statements were that they are cruel and mistreat the Jews. The 47 per cent neutral concepts were that they had a boy king, are ruled by King Carol II, who cares more for love than his crown, and they have a queen. They make carpets, dress in old-fashioned clothes, and are like Germans. Thirty-nine per cent of the reactions were "don't knows."

In 1944 again 10 per cent of the concepts were favorable. The Roumanians are neat, friendly, nice, polite, good, O.K., and gay, religious, skillful, and wear pretty clothes.

Eight per cent unfavorable responses described them as poor, odd, hard, cruel, nasty, two-timer foreigners who belong to the axis and are not liked. Fifty-seven per cent of the concepts were neutral. These were oil-fields, Rome, Roman, Roumania, the war, southeastern Europe, small country, and forests. Twenty-five per cent were "don't know."

In 1931, the 6 per cent favorable concepts for Arab were that they travel on fast horses and raise date trees. Fifty-four per cent were unfavorable responses. They travel and live on deserts, rob rich desert travelers, are dark and war-like. They have trouble with the Jews in Palestine and tried to tear down the wailing wall. Thirty-one per cent were neutral concepts. Arabs come from a hot country, use camels like we do horses, travel in caravans, speak Bohemian language, and wear white robes.

In 1944 10 per cent of the concepts were favorable. The Arabs like horses,

have veiled pin-up girls, are hospitable, generous, O.K., beautiful, strong, and enduring. They can stand heat and are brown but belong to the white race.

Forty-seven per cent of the responses were unfavorable. The Arabs live on deserts, in tents. They are dark, dirty, smell, are wild savages, unkind, bad, tricky bandits, war-like, robbers, and thieves. They are Nomadic Bedouins, wanderers, and travelers.

Thirty-six per cent neutral concepts were Arabia, ride camels, Moslems, turban, white clothes, veils, women wear masks on face, swords, mysterious, and different.

Five per cent of the concepts for Chinese were favorable in 1931. They said the Chinese adopted American ways, were first to make silk, are skillful in making Chinaware and dolls.

Fifty-eight per cent of the reactions were unfavorable. The Chinese belong to the Mongolian race, have yellow skin and slanting eyes, wear ques, and are funny looking. They are sneaky, cruel, barbarous, ugly, and have bad tempers. They are war-like, revenge murders, and kill helpful missionaries. They are smugglers, originate opium drugs, and have opium dens. Immigration of coolies is prohibited. They are different from us and give me the chills in the movies. They do laundry work. Thirty-eight per cent neutral concepts said they eat rice, eat with chop sticks, are like the Japanese, built a wall, are densely populated, have hard times and famines every winter. They are small and have small feet. They pull people in carriages. The Chinese play an important part in "Welcome Danger."

In 1944 the concepts of the Chinese show a definite trend toward more favorable attitudes. Thirty-seven per cent of the reactions were favorable. They are nice, kind, friendly, good, helpful, polite, smart, swell, and good-natured people. They are good, rugged, brave fighters and allies. They are ancient people, wonderful, hard-working, quiet, simple, and respected. They are pretty, handsome, make designs, and silk, and bright clothes. Their government is good. The Chinese are fine cooks.

Unfavorable remarks had decreased to 30 per cent. The Chinese are yellow-skinned, with slanting eyes and ques. They are poor, have no food, go without shoes, are queer, talk funny, and have funny customs. Thirty-three per cent neutral concepts said they are people from China, are small, old-fashioned, look like Japs, have dark hair, eat rice and fish. Other associations were, three hundred dollars for a meal, paper walls, huts, boats, peasants, Chinatown, Orientals, East, Madam, and my uncle is with them fighting.

Two per cent favorable responses to French-Canadian, in 1931, said they work hard. There were no unfavorable remarks. The 47 per cent neutral concepts said they were of French ancestry, live in Canada but come from France, are a mixture of people, are half French half Canadian, are mostly French and American. The "don't knows" were 51 per cent of the reactions.

In 1944 there were 18 per cent favorable concepts. The French-Canadians are brave, nice, polite, good, swell. They are allies and good fighters. They are good fishermen. The Quintuplets are French-Canadian. One per cent unfavorable response was don't like them. Thirty-two per cent neutral responses said they are a mixed race and speak two languages. They are lumbermen, traders, trappers, and fur hunters. They live in a cold country and have Canadian dogs. Other reactions were geographical names and places. "Don't know" comprised 55 per cent of the reactions.

The word Mulatto has little meaning to the children. In 1931 one per cent responses said they were friendly people and 6 per cent said they were half white and half Negro. Five per cent said they had a tan skin and 88 per cent said "don't know."

In 1944 there were no favorable concepts, 5 per cent said they were half white and half Negro. Ninety three per cent said "don't know."

In 1944 there were 42 per cent favorable concepts of South American. They are rhythmic people who like music, songs, dancing, Rumba, Tango, woo-woo, ta-ra-ra-boom. They are colorful, have pretty dancing costumes and grass skirts, wear pretty clothes, full skirts, and much jewelry. It is a nice, colorful country with pretty girls who have much fun. The people are beautiful, gay, romantic, jolly, bright, nice, good, swell, O.K., wonderful, and warm. They are like us, dress like us, are good neighbors, and make good coffee. The Panama Canal is there.

Five per cent unfavorable remarks said they have dark skin, are natives who revolt, have riots, and live in jungles and swamps. Forty-eight per cent of the concepts were neutral. They mentioned Brazil, hot climate, person from South America, and South; also Amazon, Argentina, Andes, swamps, forests, and old America.

There are 70 per cent favorable concepts of Hawaiian. The ladies wear grass skirts, they have pretty dancing girls, gay dancing people, swell dancers, and beautiful Hawaiian girls who wear sarongs. They have nice clothes, are handsome, pretty, and glamorous, musical, music lovers, and have songs, guitars, banjos, and jazz. Aloha, oh boy, they are cute, work hard, are good

swimmers, and are allies in war. In Hawaii they have pretty palm trees, flowers, pineapples, cocoanuts, dates, and sugar cane.

Five per cent unfavorable concepts said they are brown people, mysterious, uncivilized natives who do not wear much clothing. Twenty-two per cent were neutral concepts. They mentioned Pearl Harbor, from Hawaii, South Pacific, volcano, mountains, straw hats, black hair, and different language.

Seventy-two per cent of the concepts for Australia were favorable. They have many kangaroos, animals, ostriches, birds, rabbits, guinea-pigs, moose, Teddy bears, and sheep. Australia is called the country down under. It is a beautiful country with woods, trees, and plants. The people are friendly, have a warm feeling, and are nice, swell, good, O.K., jolly, gallant, and very husky. They are white people, much like Americans, citizens of England, and like the English, speak English with a nice accent, and are called Aussies. They are good fighters, helped supply the United States army, and use the boomerang. MacArthur was there.

One per cent unfavorable remarks said they were wild savages. The 23 per cent neutral concepts mentioned Australia, far away, herders, hot country, brown from sun, soldiers wear funny hats, and brother is stationed there.

D. SUMMARY

For the periods 1931 and 1944 the children showed little change in favorable concepts of Turk, Portuguese, Roumanian, Arab, and Mulatto. There were many "don't knows" for all of these peoples, except for the Arab of whom there were a large number of unfavorable concepts.

There was a large increase in favorable responses and decrease in unfavorable ones for Chinese, and also an increase in favorable concepts of French-Canadian. The children studied in 1944 expressed a high percentage of favorable concepts of South American, Hawaiian, and Australian. The data suggest that not knowing much about certain races or nationalities is accompanied by few favorable concepts of them. The children studied in 1931 and 1944 expressed similar concepts and stereotypes.

REFERENCES

1. BAYTON, J. A. The racial stereotypes of Negro college students. *J. Abn. & Soc. Psychol.*, 1941, **36**, 97-102.
2. BLAKE, R., & DENNIS, W. The development of stereotypes concerning the Negro. *J. Abn. & Soc. Psychol.*, 1943, **38**, 525-531.
3. GREGORY, W. S. A study of stereotypes thinking; affective reactions to persons as the basis for judging their nationality. *J. Soc. Psychol.*, 1941, **13**, 89-102.

4. KUHLEN, R. G. Attitudes toward enemy and allied countries. *J. Abn. & Soc. Psychol.*, 1943, **38**, 277-283.
5. MEENES, M. A comparison of racial stereotypes of 1935 and 1942. *J. Soc. Psychol.*, 1943, **17**, 327-336.
6. MELTZER, H. Children's thinking about nations and races. *J. Genet. Psychol.*, 1941, **58**, 181-199.
7. SHOENFELD, N. An experimental study of some problems relating to stereotypes. *Arch. of Psychol.*, 1942, **38**, No. 270.
8. ZELIGS, R. Racial attitudes of Jewish and non-Jewish children in the sixth grade. Unpublished Master's thesis, University of Cincinnati, 1932.
9. ———. Racial attitudes of children as expressed by their concepts of races. *Sociol. & Soc. Res.*, 1937, **21**, 361-371.
10. ———. Children's intergroup concepts and stereotypes. *J. Ed. Sociol.*, 1947, **21**, 113-126.
11. ———. Your child's racial attitudes. In *Glimpses into Child Life*. New York: Morrow, 1942.
12. ———. Trends in Children's Intergroup Attitudes. (Unpublished manuscript.)

14284 Dickens Street
Sherman Oaks, California

A NOTE ON DOMINANCE TESTS*

Department of Psychology, University of Maryland

SHERMAN ROSS¹

A. INTRODUCTION

The purpose of this note is to state a problem, to discuss some limited findings in regard to social dominance in dogs, and to raise several methodological considerations.

Dominance ordering in animal groups has become a key concept in describing the social interactions among individual members of these groups. The work of Scheldrup-Ebbe on chickens created an initial experimental interest in this area, and his work has been carried forward in many different species. A review of the recent literature of the past decade by Smith and Ross (1) shows clearly the interest of contemporary workers in this phenomena. The great emphasis placed on dominance would appear to make this kind of social interaction the most important characteristic of social activities. In the study of non-human animal behavior, it is easy to understand why this prominent and readily observable behavior should attract the attention and interest of research workers.

In view of this interest, the writer believes that several points need to be made. This paper then concerns itself with three ideas: (a) the relative reliability of data on dominance relationships, (b) the need for stressing the complexes of factors in situations in which dominance is studied, and (c) the degree of generalized and/or specialized dominance characteristics.

A most important point of view in regard to this note has been expressed by a distinguished leader in the field of comparative psychology. He is of the opinion that those readers who are already in agreement with the major theses will not need data to convince them of the validity of the arguments. Other people, however, may be convinced that current dominance tests *are* reliable. It seems then that the arguments for the points made might be jeopardized by the lack of complete data in their support. The present writer, however,

*Received in the Editorial Office on July 30, 1951.

¹This study was done at the Division of Behavior Studies, Hamilton Station, R. B. Jackson Memorial Laboratory, Bar Harbor, Maine, during the summer of 1947. The writer should like to express his thanks to his wife, Jean Goodwin Ross and to Mr. Frank H. Clark for their help in securing the data reported here. To the Laboratory for its kind support, a great debt is acknowledged.

believes that in spite of the paucity of the supporting evidence that it is important that the points be stated, and this belief provides the sole rationale for the preparation of this note.

The data to be presented serve only to illustrate the problems and are not considered to be complete or definitive. It is realized that the results based on 36 trials deal with four subjects of a particular hybrid at a given time in their development. The problem, nevertheless, is of importance and in the opinion of the writer calls for a solid study of the notions expressed herein.

B. SUBJECTS AND PROCEDURE

The results reported were gathered during the summer of 1947 at the Division of Behavior Studies, R. B. Jackson Memorial Laboratory, Bar Harbor, Maine.

1. *Subjects*

The subjects for the experiment were two male and two female Chow-Basenji puppies. The puppies were three months old at the start of the experiment which lasted for about 21 days. These puppies, their characteristics, care, and housing have been described in detail in a previous paper (2).

2. *Procedure*

Three situations for the testing of food related dominance were used: dish dominance, bone dominance, and exit dominance. Dish dominance refers to a situation in which a small dish (7" diameter, $2\frac{1}{2}$ " deep) was presented to each pair of dogs. The dish was selected so that it would lead to a space restricted feeding situation. All of the subjects had been previously fed from a much larger dish. The bone dominance situation refers to the situation where a beef bone with some scraps of meat was thrown between a pair of dogs. The exit dominance refers to the situation where the dogs were permitted to enter a small box in which there was a small corner compartment with a dish of food. When the dog ate from the dish, he was exposed to attack from behind by the other dog in the pair.

All of the tests were made at regular feeding times, either morning or afternoon. No dog took part in more than one dominance test on a single day. Systematic pairings were made and repeated. Each session lasted 10 minutes.

During each test session, the following items were observed and recorded: biting, growling, eating, urination, and defecation. A dog was considered to be dominant when he secured and defended the food pan or bone and ate the food.

TABLE 1
THE DOMINANT DOG OF EACH PAIR FOR EACH TEST SITUATION

Pairing	Dish dominance		Bone dominance		Exit dominance	
	Trial 1	Trial 2	Trial 1	Trial 2	Trial 1	Trial 2
A × B	B	—	A	A	B	B
A × C	C	—	A	A	C	C
A × D	D	A	D	D	D	D
B × C	—*	—	B	C	C	C
B × D	D	D	B	B	B	B
C × D	D	—	C	D	C	C

*Indicates that no clear dominance was observed. In most cases both dogs ate the food in the dish.

C. RESULTS AND DISCUSSION

The basic results of the dominance tests, including repeat tests on the same pairs, are shown in Table 1.

The table shows that an exact duplication of the dominance relationship does not appear for either the dish or the bone dominance situations. There is only one case ($B \times D$) in which the same results appeared on Trial 2 for the dish dominance, and four cases of the six pairings where the same puppy was dominant in the repeat trial of the bone dominance. In the case of the exit dominance tests, the same results appeared on repetition.

Thus, it would seem reasonable to conclude that variability is a major factor in dealing with such data. It is difficult to know how to interpret the results of a test such as dish dominance, when on repeating the test, a diversity of results are found.

The bone dominance tests are not surprising in terms of difficulty of doing such tests in a manner in which they can be repeated. The problem of tossing a bone equidistant between a pair of dogs, which may be responding to the bone, the experimenter, or many other things, is one for a skilled hurler.

That the exit dominance situation should have given such reproducible results is interesting. This may be interpreted in terms of the relative restriction of the situation. The variability was more limited. When the dogs were released to the box, one would run in, eat the food quickly, accept the attack, and leave. The fights in these hybrids were of particular interest because of their apparent vigor and extreme quiet. The Chow-Basenji hybrids inherited the "barkless" qualities of the Basenji father.

The writer recognizes several features of the present study that affect the results obtained. Among such factors may be mentioned that the subjects were young, that the observations were limited, that variability is to be expected and that variable strengths of incentives might be present.

To summarize these points it may be said that (a) it is of the greatest importance to know about the reliability of the measures in whatever situation is being used as measures of dominance, and (b) that if a dog is dominant in Situation A it does not necessarily follow that he is also dominant in Situation B, which may or may not deal with the same kind of goal. The question may be raised whether or not one should deal with dominance as a general trait to describe a social interaction, or as a specific kind of behavior elicited in a specific situation under certain conditions at a given point of time.

REFERENCES

1. SMITH, W. I., & ROSS, S. The social behavior of vertebrates: Review of the literature (1939-1950). *Psychol. Bull.*, 1952 (in press).
2. ROSS, S., & ROSS, J. G. Social facilitation of feeding behavior in dogs: I. Group and solitary feeding. *J. Genet. Psychol.*, 1949, **74**, 97-108.

Department of Psychology
University of Maryland
College Park, Maryland

BOOKS

The Journal of Genetic Psychology, the *Journal of General Psychology*, and the *Journal of Social Psychology*, will buy competent reviews at not less than \$2 per printed page and not more than \$3 per printed page, but not more than \$15.00 for a single review.

Conditions. Only those books that are listed below in this section are eligible for such reviews. In general, any book so listed contains one or more of the following traits: (a) Makes an important theoretical contribution; (b) consists largely of original experimental research; (c) has a creative or revolutionary influence in some special field or the entire field of psychology; (d) presents important techniques.

The books are listed approximately in order of receipt, and cover a period of not more than three years. A reviewer must possess the Ph.D. degree or its equal in training and experience.

Procedure. If among the books listed below there is one that seems important to you, you are invited to write a review of that book. It is not necessary to make arrangements with the Editor. Just send in your review. It does not matter if the book in question has been reviewed before.

1952

- STONE, C. P., *Ed.* Annual Review of Psychology, Vol. 3. Stanford, Calif.: Annual Reviews, 1952. Pp. 462.
- HOCH, P. H. & ZUBIN, J., *Eds.* Relation of Psychological Tests to Psychiatry. New York: Grune & Stratton, 1952. Pp. 301.
- CATTELL, R. B. Factor Analysis. New York: Harper, 1952. Pp. 462.
- FOERSTER, H. VON, *Ed.* Cybernetics. New York: Macy Foundation, 1952. Pp. 240.
- BENDER, L. Child Psychiatric Techniques. Springfield, Ill.: 1952. Pp. 335.
- METTLER, F. A., *Ed.* Psychosurgical Problems. New York: Blakiston, 1952. Pp. 357.
- BENDA, C. E. Developmental Disorders of Mentation and Cerebral Palsies. New York: Grune & Stratton, 1952. Pp. 565.
- KINGET, G. M. The Drawing-Completion Test. New York: Grune & Stratton, 1952. Pp. 238.
- WOLFF, C. The Hand in Psychological Diagnosis. New York: Philosophical Library, 1952. Pp. 218.
- WOLFF, W. The Dream: Mirror of Conscience. New York: Grune & Stratton, 1952. Pp. 348.
- GORLOW, L., HOCH, E. L., & TELSCHOW, E. F. The Nature of Nondirective Group Psychotherapy. New York: Teach. Coll., Columbia Univ., 1952. Pp. 143.
- AUSUBEL, D. P. Ego Development and the Personality Disorders. New York: Grune & Stratton, 1952. Pp. 564.
- BYCHOWSKI, G., & DESPERT, J. L. Specialized Techniques in Psychotherapy. New York: Basic Books, 1952. Pp. 371.
- WHITE, R. W. Lives in Progress: A Study of the Natural Growth of Personality. New York: Dryden, 1952. Pp. 376.
- KARN, H. W., & GILMER, B. VON H. Readings in Industrial and Business Psychology. New York: McGraw-Hill, 1952. Pp. 476.
- ROBACK, A. A. History of American Psychology. New York: Library Pub., 1952. Pp. 426.
- KROEBER, A. L. The Nature of Culture. Chicago: Univ. Chicago Press, 1952. Pp. 437.
- FERGUSON, L. W. Personality Measurement. New York: McGraw-Hill, 1952. Pp. 457.

- BECK, S. J. Rorschach's Test: III. Advances in Interpretation. New York: Grune & Stratton, 1952. Pp. 301.
- VINACKE, W. E. The Psychology of Thinking. New York: McGraw-Hill, 1952. Pp. 392.
- KUHLEN, R. G., & THOMPSON, G. G., *Eds.* Psychological Studies of Human Development. New York: Appleton, Century, Crofts, 1952. Pp. 533.
- SNYDER, F. W., & PRONKO, N. H. Vision with Spatial Inversion. Wichita: Univ. Wichita Press, 1952. Pp. 144.
- JERSILD, A. T. In Search of Self. New York: Teach. Coll., Columbia Univ., 1952. Pp. 141.
- HIRSH, O. J. The Measurement of Hearing. New York: McGraw-Hill, 1952. Pp. 364.
- ASHBY, W. R. Design for a Brain. New York: Wiley, 1952. Pp. 260.
- SHAFFER, G. W. Fundamental Concepts in Clinical Psychology. New York: McGraw-Hill, 1952. Pp. 540.
- HAYEK, F. A. The Sensory Order. Chicago: Univ. Chicago Press, 1952. Pp. 209.
- AMES, L. B., LEARNED, J., METRAUX, R. W., & WALKER, R. N. Child Rorschach Responses. New York: Harper, 1952. Pp. 310.

1953

- KLUCKHORN, C., MURRAY, H. A., & SCHNEIDER, D. M., *Eds.* Personality. New York: Knopf, 1953. Pp. 701.
- SKINNER, B. F. Science and Human Behavior. New York: Macmillan, 1953. Pp. 461.
- STONE, C. P., & TAYLOR, D. W., *Eds.* Annual Review of Psychology. Stanford: Annual Reviews, 1953. Pp. 485.
- BURGESS, E. W., & WALLIN, P. Engagement and Marriage. New York: Lippincott, 1953. Pp. 819.
- POWDERMAKER, F. B., & FRANK, J. D. Group Psychotherapy. Cambridge: Harvard Univ. Press, 1953. Pp. 615.
- PHILLIPS, L., & SMITH, J. G. Rorschach Interpretation: Advanced Technique. New York: Grune & Stratton, 1953. Pp. 385.
- MURPHY, G. In the Minds of Men. New York: Basic Books, 1953. Pp. 306.
- BLUM, G. S. Psychoanalytic Theories of Personality. New York: McGraw-Hill, 1953. Pp. 219.
- BROWER, D., & ABT, L. E., *Eds.* Progress in Clinical Psychology. New York: Grune & Stratton, 1953. Pp. 564.
- GELDARD, F. A. The Human Senses. New York: Wiley, 1953. Pp. 365.
- CABOT, H., & KAHL, J. A. Human Relations. Cambridge: Harvard Univ. Press, 1953. Pp. 333. & 273.
- SHERIF, M., & SHERIF, C. W. Groups in Harmony and Tension. New York: Harper, 1953. Pp. 316.
- WHITING, J. W. M., & CHILD, I. L. Child Training and Personality. New Haven, Conn.: Yale Univ. Press, 1953. Pp. 353.

CRITICAL REVIEWS OF RECENT BOOKS

The Journal of Genetic Psychology, 1953, **83**, 185-194.

(Cattell, R. B. *Factor Analysis: An Introduction and Manual for the Psychologist and Social Scientist*. New York: Harper, 1952. Pp. 462.)

REVIEWED BY A. S. C. EHRENBERG

A. INTRODUCTION

In this new account of factor analysis, Cattell's major aim has been to discuss the subject more simply than has been done in previous treatises. The general tenor of the beliefs expressed and of the computations given is Thurstonian. Apart from describing the aims and the place of factor analysis in scientific methodology, Cattell's main topics are the well-known centroid solution and its subsequent rotation, particularly by applying the principle of simple structure. The nature of O-, P-, Q-, and R-techniques and problems of the design of factor analytical experiments are also touched upon, and an appendix for certain electronic computations, a glossary, a large bibliography, and indices are given. The book may therefore be of especial value to those who require a not too technical account of the computational procedures of multiple factor analysis, and also of what is usually said about them.

Factor analysis, besides being rather a controversial subject, is an unusual one. It is essentially a theoretical piece of mathematics, but as a statistical method of handling empirical data, it differs from orthodox statistics. Thus it is often suggested that the method possesses some *a priori* power to discover "causal" scientific relationships, though actual claims for its methodological validity rest less on theoretical proofs or arguments than on empirical findings.

Consequently, there exist a variety of points of view from which a text-book such as this might be evaluated. Simply to consider its clarity of exposition and to note new matter, or just to judge the book statistically up to date because analysis of variance is mentioned in its opening sentence, seems hardly satisfactory. Again, to attempt deep comment on the general principles of factor analysis might be thought out of place in reviewing an introductory text, a manual. Therefore, apart from considering very briefly the degree of scientific and statistical sophistication which has been brought

to the task, we cannot perhaps do better than examine some of the more fundamental computational procedures put forward. Have these actually been shown to satisfy the theoretical principles of the factor analytic approach, whatever the latter's merits may be?

It will probably be sufficient for a book review to restrict ourselves to quite a small selection of such technical points, and we shall consider a few facets of the centroid methods, of factor rotation, and of factor identification. Has it been proved, we must ask (either in the book or, by reference, elsewhere) that the methods given do whatever they are said to do?

B. SCIENCE, CATTELL AND STATISTICS

In discussing the place of factor analysis in scientific methodology, Cattell repeatedly stresses that "the factor analyst is suspicious" of choosing the important variables *a priori*," and that "he would like to find the real independent factors . . . before entering into experiment with them" (p. 15). Factor analysis contributes "to scientific understanding of what basic influences are operative" (p. 20). Cattell explicitly raises factor analysis above other statistical methods of analysis: multiple and partial correlations for example are contrasted with it as "scarcely tools of investigation" and are said to "contribute little or nothing . . . in terms of scientific concepts" (p. 20).

As already mentioned, it is not our purpose in this review to comment on the possibilities of factor analysis to reveal "wholes" or "functional unities" which are "causal" and "scientifically meaningful"—one can discuss the computations without that.¹ But it must be mentioned that Cattell's use of such terms as "causal" and "functional unity" seems none too precise. One may indeed wonder what purpose extended discussion based on such terms has in this book, since on the one hand "instances when the factor corresponds to an organic unity are admittedly hard to find" (p. 318), whilst "as to causal sequences, it is probable—though as yet it rests only upon a *a priori* argument—that factor analysis can directly throw light on the sequence" (p. 362). In other words, after 50 years of empirical factor analysing we are told that the mere probability even that the method might unravel what Cattell considers causal processes rests only on a *a priori* (but unstated) grounds.

¹Professor Cattell has kindly queried the large number of words in quotation marks. The marks are supposed to indicate (i) that I am quoting and (ii) that I do not necessarily take responsibility for the meaning attached to the words in the present kind of context.

Factor analysis may, or must, therefore be regarded as just another statistical technique, like calculating averages, standard deviations, and so on, whose first requirements are theoretical consistency and understanding, and whose scientific usefulness, where it exists, can only be established empirically. However, the book itself can hardly be thought to adhere to a down-to-earth, descriptive statistical outlook. For example, the old dream that correlation reads causation is resuscitated in the case of factor loadings (p. 362): "it would seem that in general the variables highly loaded in a factor are likely to be the causes of those which are less loaded, or, at least that the most highly loaded measure—the factor itself—is causal to the variables that are loaded in it."

In general, there can be little question that Cattell's statistical pronouncements—whether right or wrong—lie well outside the statistical canon. Thus for the comparison of simple *averages*, we find it stated (pp. 10 and 364) that an analysis of variance (i.e., an investigation of the differences between averages) indicates only the existence of a relationship, but not its degree or strength. But is it not just the "degree" of the relationship which is expressed by the differences between the averages and by the size of the variances?

Turning in particular to the treatment of correlations, the raw material of factor analysis, we find that the product-moment correlation coefficient is considered "best" for factor analysis, that no hesitation is shown in saying that *normality* is not "required," and that even *J*-shaped distributions are thought not to "invalidate" the use of the coefficient. This might all be acceptable if only one were told in what sense these things were "best," nor "required" and not "invalid"—as it is, the assertions are not only unproven but logically unprovable. However, when Cattell supports Thurstone even so far as "to select the subjects so that their attributes are as diverse as possible in the domain to be studied" (p. 354), a statistical rather than purely logical criticism arises. Statistical correlations *only* describe the relative frequency with which certain of the combinations of observed attributes occur oftener together than others, and they become almost meaningless (i.e., no deductions can be made from them) if these frequencies are deliberately manipulated. The point is not, of course, that any such selection of extremes is wrong, but that simple measures of statistical correlation (and *ipso facto* technical factor analysis) are inadequate, to say the least, for whatever the factor analyst's purpose might then be.

C. THE INITIAL SOLUTION

1. *The Simple Centroid Method*

Technically, factor analysis resolves itself upon some initial technique of extracting "factors" from a given array of correlations, although Cattell believes that such factors must subsequently be "rotated" in order to be useful. The preferred initial solution given in the book is the well-known *centroid* one; i.e., one adds up each column in the correlation matrix and divides each such sum by the square root of the sum of *all* the correlations. This gives the so-called factor saturation or factor loadings (compare the numerical example below).

After one such centroid factor has been extracted and before the next one can be obtained, some of the residual correlations have to be "reflected," i.e., the signs of a number of the residuals have to be changed. (Sometimes even some of the original correlations are reflected.) Obviously, this process of reflection can be carried out in a variety of different ways, each yielding a different set of factors. To the normal scientist or statistician such a lack of uniqueness would of course not matter greatly (he knows that any given data can be represented variously) as long as the implications of the different ways of reflecting are made clear, so that he can choose that which seems most useful, and in any case have some idea of what he is doing to his data.

However, no relevant information is given in the book. What *are* the different methods that could be adopted and what would their consequences be? Even the effects and implications of the particular computational process which is actually adopted are made almost impossibly difficult to grasp: "For as Diagram 8 shows, we need to bring all the test vectors on one side . . ." (p. 55). But Diagram 8 shows nothing of the kind; it illustrates what things would look like if one *did* "bring all test vectors on one side," but it certainly does not give any *reason* for doing so. What is more, even just to illustrate the meaning (as opposed to the need) of the given mode of reflection, the diagram would seem to beg the question. Two "factors" are represented with their loadings, but it is these very loadings which hinge on the particular process of reflection adopted.

2. *The Choice of Communalities and the Factor Model*

So far we have passed over one dilemma in the centroid method (or any other kinds of factor analysis), namely that there are no "diagonal" terms in the original matrix of correlations. By definition, any (orthogonal) factor solution is required to be such that, for any two variables, the products of

the respective factor saturations add to the original correlation (compare the numerical example below). Clearly, it is possible to multiply the factor saturations of a single variable also with themselves, which would result in numbers that could be conveniently entered in the diagonal of the matrix of reproduced correlations. It then seems to be felt desirable to enter some such diagonal values already in the original correlation matrix, before one begins the actual factoring. How is this problem treated in Cattell's book?

In the book, the use of diagonals seems to be taken for granted, so that the first question to arise is how actually to choose their values. Obviously, to obtain factor saturations which could possibly be interpreted as correlation coefficients, the diagonal values must be either just unity, or so-called *communalities* lying between zero and plus one. We find that the first choice, that of unit diagonals, is dismissed simply because "tests do *not* have nothing but common factors; they certainly have some specifics" (p. 157). However, the usual factor model of so-called *common* and *specific* factors arises directly, by logical necessity, out of communalities which are not unity, and similarly, if non-unity communalities are chosen, one *must* set up a model of common and specific factors. It seems therefore impossible to use the model to *justify* the choice of communalities, or vice versa.

Anyway, whatever his reasons, Cattell does embrace the traditional model of "multiple" factors, with common, specific, and error factors, and with non-unity communalities. Two further questions now arise. What principles are given to guide one in the choice of actual diagonal values, and what are the known consequences?

As to the principles, Cattell denies (pp. 156, 295) that in choosing communalities he is minimising this (e.g., the number of factors) or maximising that (e.g., factor loadings), or rather, he denies that he is intentionally trying to do any such thing, for he does not attempt to prove that he is not actually doing so. Further, although Cattell reports Thurstone's argument about matching "the true rank as indicated by the best fit with the existing correlations, i.e., the off-diagonal r 's" (p. 295), this gives (according to Thurstone, at least) simply the minimum rank (or minimum number of factors), which Cattell has already rejected. It seems to follow that Cattell not only fails to prove that his choice of numerical communalities satisfies some principle or other, but that he has not even set up any principles, whether acceptable or not, to be satisfied.

What, then, are the main consequences of following one of the several purely computational procedures given by Cattell for choosing communalities? In view of the lack of theoretical principles to judge one of these

procedures in some sense *best*, or even just *better* than any of the infinite number of possible choices of communalities which are *not* given, it might be supposed that no very general conclusions could be drawn. But that seems not altogether to be the case.

For all factor analyses with non-unity communalities, no way of calculating any score on any factor for any individual is known. Cattell here rather gives the wrong impression, as for example when he describes centroid factors as "always estimated with a margin of error" (p. 132). Not only are centroids based on *unit* diagonals quite determinate, but it is not a question of "error" in the usual statistical sense at all. The important point *here* is probably to emphasise that no method (which works) has ever been given for calculating multiple factor scores; but one must also note the absence of any sort of proof that such methods could, in fact, ever be found. Actually, multiple factor scores do not seem to exist, and if we are to "give every individual his new apparel of factor measurements to replace the motley rags of his original numerous scores" (p. 73), these measurements can but be "the emperor's new clothes" (1).

3. *Some Computational Errors*

Ignoring any of the problems that would arise from dealing with correlations based on just *sample* data, the centroid method would appear to work. Successive "orthogonal" factors are put through the "centroids" of the original and of the various residual correlations respectively, with factor loadings which reproduce the original correlations. We may not know quite why this is done, but since no particular attempt is made by Cattell in his book to interpret any such initial solution, it is perhaps unnecessary to enquire what the meaning or usefulness of "centroids" might be. However, before going on to consider the "rotation" of this initial solution, one or two further computational features of the centroid method deserve comment.

The first concerns the advocacy, following Thurstone, of introducing *new* communalities for each successive factor in the simple centroid method just discussed (p. 56). Nothing much seems known about the consequences of this, but it seems unlikely that the factors so obtained would be "orthogonal" as is required. This is of course just a guess; but then factor analysts as such do not seem to have proved that even the proper centroid factors are, in fact, orthogonal (although it seems to be true there), and guesswork must be excusable.

The second point concerns various other computational centroid pro-

cedures reported from Thurstone, so-called "clustering" methods. These are apparently thought to be alternatives to the simple centroid method (otherwise there would seem to be little point to them), but no attempt is made to prove this. Actually, the new methods do not appear to give the same kind of solution at all. For example, they would not necessarily give the same number of "factors."

Take as an example two tests made up of one "common" factor, with saturations .4 and .3 respectively. This gives a 2 x 2 correlation matrix

$$\begin{array}{cc} .16 & .12 \\ .12 & .09 \\ \hline .28 & .21, \quad \sqrt{.49} = .70, \end{array}$$

and just a single centroid factor with saturations $.28/.7 = .4$ and $.21/.7 = .3$ (as postulated). In the first of the clustering methods, the "grouping" one (p. 172) for example, one would perform a centroid kind of solution on a group of the tests, in this case for instance just the first one. This gives a first "factor" with "saturations" $.16/\sqrt{.28}$ and $.12/\sqrt{.28}$, and, of course, residuals. In other words, there would be a further "common factor" and the method can hardly be said to work.

D. FACTOR ROTATION

As already mentioned, Cattell believes that factors such as are obtained in a centroid solution have to be "rotated" in order to be "meaningful." Of various methods of rotation which have been suggested, Cattell only accepts Thurstone's principle of "simple structure"; amongst those dismissed is Eysenck's "criterion analysis," the computational procedure of which has however (so Dr. Eysenck tells me) been misrepresented (p. 250).

The simple structure principle, we are told, "is fortunately able to meet the demands of practically all experimental designs" (p. 251). Thurstone's general problem seems to be to see if each of the given test variables can be described separately by fewer common factors than in the initial factor solution; the suggested *solution* consists, roughly, of rotating until enough factor loadings are zero. The appeal of all this as a universal scientific panacea shall not concern us here, but rather its working in practice.

The first thing we find is that the subject becomes vastly more complex largely, as Cattell rightly stresses, because of the departure from "orthogonal" factors to "oblique" ones. For these factors one "must accept it as a mathematical truth (if he does not have insight in this realm) that all calculations which have to do with factors really define the factors as the

line of intersection *for all hyperplanes other than that of the factor itself*" (p. 218). Hyperplanes, we note, are certain spaces—one for each factor, apparently—of one less dimension than the total number of (initially extracted) factors. However, we find that if we turn even to the simplest examples, this geometrical picture is by no means obvious. Thus with just two initial factors (one-factor solutions seem to be out of it altogether) the hyperplanes are lines, which can only intersect in a point (the origin here), and it is not clear what the "factors" are.

As for the algebra of oblique rotations, we are told that a factor loading does *not* stand for the correlation of the test and factor in question. In fact, it would seem that this correlation—as far as it exists—can take any value whatever. In particular, a zero entry in the final simple structure solution does not imply that the test and factor are uncorrelated. It seems difficult therefore not to agree with the suggestion (p. 222) that "this anomaly should stimulate us as a practical and theoretical issue at least to weigh the advisability of"—in short—not using these particular "factors" at all.

What then are the properties of the factor analytic solution obtained by the application of this simple structure principle, apart from the one that the factors obtained do not, by general agreement, satisfy the principle? Apparently, about the only other property suggested is that such an analysis produces stable results of which the short cut (?) of ordinary correlation theory are not capable (p. 385). Such an attractive property of the rotated solution is implied, one presumes, when unrotated factors are said to lack "invariance," and are therefore deemed "psychologically and statistically meaningless" (p. 249). At one stage (p. 305) we are actually told quite explicitly that "ample reasons have been given earlier (p. 123) for believing that invariance will be attainable only when the results . . . are rotated to simple structure."

However, when we find that no reasons at all, whether acceptable or not, are given on page 123, that no sort of case has been made out even for the desirability of "invariance" itself, and that, as regards proof, "it cannot at the moment be adequately proved or disproved" (p. 249) whether or not unrotated factors give invariance as well as does the simple structure method, it is not surprising that the matter should be summarised by the sentiment (p. 305) that "researchers are entitled to believe in the possibilities of obtaining invariance with *any* system. . . ."

E. FACTOR IDENTIFICATION

Having factor analysed a set of correlations, the factors obtained need somehow to be identified and interpreted. This problem does not appear to have been systematically treated so far, and the absence of anything like a

definite solution may well be a major reason for the orthodox statistician's tendency to disregard factor analysis for practical use. Cattell seems to break new ground here, although he is perhaps rather prescient in discussing such a problem in a chapter headed "The effects of errors."

In multiple factor analysis (i.e., with non-unity communalities) no factor scores are, as we have already noticed, defined, and factor interpretation therefore proceeds traditionally by the inspection of the factor loadings, these being apparently considered as correlations of the test scores with the factors, whatever the latter may precisely be. Thus, a factor is "identified" by the test variables which are highly loaded in it. Now, as with the topics already discussed, we may find the mere technical (and unresolved) difficulties of this procedure sufficiently daunting, but in this case we should perhaps also note that, as far as one can judge, factor identification by its test loadings would seem to beg the question in its general aspect even.

Factor analysis is carried out to try to make some sense of the original test variables, and one cannot now reverse the process, not, that is, without considerably more awareness than is usually shown of the circular argument involved. Consider a test variable which for some given data correlates well with some acceptable criterion of examination success, clinical rating, political affiliation, or generally some X-ism. The test would then usually be said to measure X-ism under the given conditions. But does the test necessarily measure X-ism in any other set of data? Obviously not—whether or not it does so is just the sort of thing one would like to be able to find out.

Is it then reasonable to consider two factors "to measure the same thing" just because they have in *different* sets of data the same loadings with such a test, or group of tests, about which we know so little?

This point may seem either rather obvious, or, in contrast, controversial. Perhaps it is safer, therefore, to try to look at the *technical* difficulties if we are to evaluate this process of factor identification objectively. We find that two variables may have correlations as high as .7 with the same test, but still be completely uncorrelated, which is not very promising if we bear in mind the size of factor loadings so often reported and the suggestions or claims that two factors from different analyses are the same because of some such similarity in their loadings. Even in a single factor analysis it is possible for two factors to have identical, or very similar, saturations on *all* the tests, and yet to be quite uncorrelated. Cattell seems to be aware of this (pp. 285 and 306), despite his initial standpoint that there can be no question of matching two factors from one analysis (p. 304). This, all in the negative, is about as much as is known about judging factors "the same" because of their similarity of loadings.

Another procedure for identifying factors which is given in the book, "matching by elimination" (p. 309), probably needs no explicit comment here: "If all other factors have been matched, in a study with the same set of variables, the remaining important factor on one side is likely to be the same as the remaining substantial factor on the other."

F. CONCLUSIONS

Cattell's *Factor Analysis* is intended as an introduction and manual, and rather than consider the merits of the general factor analytic principles, we have tried in this review to examine what some of those practical procedures for handling data which are given in the book have been shown to do. In the centroid analysis, for example, it would appear that fundamental steps like "reflecting" and "choosing communalities" have not only not been said to satisfy any important general principles, but that no such principles have in fact been set up. In the absence also of any *a posteriori* deductions, the centroid solution presented might therefore be judged rather meaningless, and inadequate for inclusion in a laboratory manual.

Centroid factors apparently have to be "rotated" by applying the principle of "simple structure." However, the only proven property of the computational solution put forward is that it does not actually satisfy the principle of simple structure. As far as the "factors" are concerned, it does not therefore seem to differ from any other possible solution in any material effect.

Any factors, once obtained, must be interpreted. In multiple factor analysis with non-unity communalities no way of finding appropriate factor scores is known, and the only generally used criterion of interpretation is to consider the variables most highly loaded in the factor in question. Nothing much seems to be known about this procedure, except that even factors with identical patterns of saturation can be completely orthogonal, i.e., not the same.

Whilst therefore Cattell may have succeeded in his aim to produce a simpler account of factor analysis, his subject matter appears to be largely compounded of unverified speculation and rather unsuccessful guess work, and this he has not made sufficiently explicit.

REFERENCE

1. ANDERSEN, H. C. *Fairy Tales*. Oxford: Oxford Univ. Press, 1944.

Institute of Psychiatry
Maudsley Hospital
Denmark Hill
London, S.E. 5, England

BOOKS RECENTLY RECEIVED

(There will always be two pages of book titles, listed in the order of receipt, i.e., the most recently received books will be found at the end of the list.)

- SHOSTROM, E. L., & BRAMMER, L. M. *The Dynamics of the Counseling Process*. New York: McGraw-Hill, 1952. Pp. 213.
- KORNITZER, M. *Child Adoption in the Modern World*. New York: Philosophical Library, 1952. Pp. 403.
- RAVEN, J. C. *Human Nature*. London: Lewis, 1952. Pp. 226
- MONTAGU, A. *Darwin: Competition and Cooperation*. New York: Schuman, 1952. Pp. 148.
- BARLOW, F. *Mental Prodigies*. New York: Philosophical Library, 1952. Pp. 256.
- JERSILD, A. T. *In Search of the Self*. New York: Teach. Coll., Columbia Univ., 1952. Pp. 141.
- SNYDER, F. W., & PRONKO, N. H. *Vision with Spatial Inversion*. Wichita: Univ. Wichita Press, 1952. Pp. 144.
- EISLER, R. *Man into Wolf*. New York: Philosophical Library, 1952. Pp. 286.
- GUILFORD, J. P. *General Psychology*. New York: Van Nostrand, 1952. Pp. 587.
- KUHLEN, R. G., & THOMPSON, G. G. *Psychological Studies of Human Development*. New York: Appleton, Century, Crofts, 1952. Pp. 533.
- VINACKE, W. E. *The Psychology of Thinking*. New York: McGraw-Hill, 1952. Pp. 392.
- FERGUSON, L. W. *Personality Measurement*. New York: McGraw-Hill, 1952. Pp. 457.
- SHOCK, N. W., *Ed. Problems of Aging*. New York: Macy, 1952. Pp. 138.
- McKILLIP, A. S. *The Relationship Between the Reader's Attitude and Certain Types of Reading Response*. New York: Teach. Coll., Columbia Univ., 1952. Pp. 101.
- BECK, S. J. *Rorschach's Test: III. Advances in Interpretation*. New York: Grune & Stratton, 1952. Pp. 301.
- SMILEY, M. B. *Intergroup Education and the American College*. New York: Teach. Coll., Columbia Univ., 1952. Pp. 212.
- GRIFFITHS, W. *Behavior Difficulties of Children*. Minneapolis: Univ. Minnesota Press, 1952. Pp. 116.
- KROEBER, A. L. *The Nature of Culture*. Chicago: Univ. Chicago Press, 1952. Pp. 437.
- HIRSH, J. *Alcohol Education*. New York: Schuman, 1952. Pp. 107.
- ASHBY, W. R. *Design for a Brain*. New York: Wiley, 1952. Pp. 260.
- HIRSH, I. J. *The Measurement of Hearing*. New York: McGraw-Hill, 1952. Pp. 364.
- LEVINSON, A. *The Mentally Retarded Child*. New York: Day, 1952. Pp. 190.
- PERLS, F., HEFFERLINE, R. F., & GOODMAN, P. *Gestalt Therapy*. New York: Julian, 1952. Pp. 466.
- SHAFFER, G. W., & LAZARUS, R. S. *Fundamental Concepts in Clinical Psychology*. New York: McGraw-Hill, 1952. Pp. 540.
- PALMER, H. *The Philosophy of Psychiatry*. New York: Philosophical Library, 1952. Pp. 70.
- HAYEK, F. A. *The Sensory Order*. Chicago: Univ. Press, 1952. Pp. 209.
- POWYS, J. C. *The Inmates*. New York: Philosophical Library, 1952. Pp. 318.
- BERNARD, H. W. *Mental Hygiene for Classroom Teachers*. New York: McGraw-Hill, 1952. Pp. 472.
- WOLFLE, D., *et al.* *Improving Undergraduate Instruction in Psychology*. New York: Macmillan, 1952. Pp. 60.
- JUNGE, W. *Bolahun: An African Adventure*. New York: Putnam, 1952. Pp. 248.
- ITTELSON, W. H. *The Ames Demonstrations in Perception*. Princeton: Princeton Univ. Press, 1952. Pp. 88.
- MUNDT, E. *Art, Form, and Civilization*. Berkeley: Univ. California Press, 1952. Pp. 246.
- SZONDI, L. *Experimental Diagnostics of Drives*. (Trans. by G. Aull.) New York: Grune & Stratton, 1952. Pp. 220.

- DEESE, J. *The Psychology of Learning*. New York: McGraw-Hill, 1952. Pp. 398.
- DE GRAZIA, S. *Errors of Psychotherapy*. New York: Doubleday, 1952. Pp. 288.
- BISNO, H. *The Philosophy of Social Work*. Washington: Public Affairs Press, 1952. Pp. 143.
- BELLAK, L., *Ed.* *Psychology of Physical Illness*. New York: Grune & Stratton, 1952. Pp. 243.
- HALL, B. A., *et al.* *Psychiatric Aide Education*. New York: Grune & Stratton, 1952. Pp. 168.
- BROWER, D., & ABT, L. E., *Eds.* *Progress in Clinical Psychology*. New York: Grune & Stratton, 1952. Pp. 328. (Vol. 1, Sec. 1.)
- LINDNER, R. *Prescription for Rebellion*. New York: Rinehart, 1952. Pp. 305.
- STONE, C. P., & TAYLOR, D. W., *Eds.* *Annual Review of Psychology*. Stanford, Calif.: Annual Reviews, 1953. Pp. 485.
- MOUSTAKAS, C. E. *Children in Play Therapy*. New York: McGraw-Hill, 1953. Pp. 218.
- NISBET, R. A. *The Quest for Community*. New York: Oxford Univ. Press, 1953. Pp. 303.
- ABRAMSON, H. A., *Ed.* *Problems of Consciousness*. New York: Macy, 1952. Pp. 156.
- SKINNER, B. F. *Science and Human Behavior*. New York: Macmillan, 1953. Pp. 461.
- BARNETT, H. G. *Innovation*. New York: McGraw-Hill, 1953. Pp. 462.
- BOSSARD, J. H. S., *et al.*, *Eds.* *Introduction to Sociology*. Harrisburg: Stackpole, 1952. Pp. 666.
- PIAGET, J. *The Origins of Intelligence in Children*. New York: International Univ. Press, 1952. Pp. 419.
- GLUCKHORN, C., MURRAY, H. A., & SCHNEIDER, D. M., *Eds.* *Personality*. Cambridge: Harvard Univ. Press, 1953. Pp. 701.
- THOMPSON, C. B., & SILL, A. P. *Our Common Neurosis*. New York: Exposition, 1952. Pp. 210.
- FEDERN, P. (Edited by Weiss, E.) *Ego Psychology and the Psychoses*. New York: Basic Books, 1953. Pp. 375.
- DAVIDSON, A., & FAY, J. *Phantasy in Childhood*. New York: Philosophical Library, 1953. Pp. 188.
- WOLFF, W., & PRECKER, J. A., *Eds.* *Success in Psychotherapy*. New York: Grune & Stratton, 1952. Pp. 196.
- WALTERS, A., & O'HARA, K. *Persons and Personality*. New York: Appleton-Century-Crofts, 1953. Pp. 677.
- POSTON, R. W. *Democracy Is You*. New York: Harper, 1953. Pp. 312.
- KROEBER, A. L., & KLUCKHORN, C. *Culture*. Cambridge: Peabody, 1952. Pp. 223.
- BROWER, D., & ABT, L. E. *Progress in Clinical Psychology*. New York: Grune & Stratton, 1952. Pp. 564.
- TRENAMAN, J. *Out of Step*. New York: Philosophical Library, 1952. Pp. 217.
- HOYLES, J. A. *The Treatment of the Young Delinquent*. New York: Philosophical Library, 1952. Pp. 274.

\$14.00 per year
\$10.00 per volume
Single Numbers \$5.00

QUARTERLY
Two volumes per year

December, 1953
Volume 83, Second Half

Founded by G. Stanley Hall in 1891, reorganized by Carl Murchison in 1925

THE PEDAGOGICAL SEMINARY AND

JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior,
and Comparative Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under

the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

DECEMBER, 1953

(Manuscripts are printed in the order of final acceptance)

Children's conceptions of some bodily functions	199
By MARIA H. NAGY	
Children's birth theories	217
By MARIA H. NAGY	
The representation of "germs" by children	227
By MARIA H. NAGY	
W.I.S.C. patterns of a selective sample of bilingual school children	241
By GRACE T. ALTUS	
Intelligence and aging	249
By RAYMOND J. CORSINI AND KATHERINE K. FASSETT	

(OVER)

Copyright, 1953, by The Journal Press
Provincetown, Massachusetts

Entered as second-class matter May 11, 1937, at the post-office at Provincetown,
Mass., under the act of March 3, 1879

The dominant method of imagery in blind as compared to sighted adolescents By T. F. SCHLAEGEL, JR.	265
Studies in delinquency: III. An individual test By F. C. ZAKOLSKI	279
Compensatory behavior as completion of the sensory-motor cycle: An interpretative hypothesis By HENRY GUZE	293
SHORT ARTICLES AND NOTES	303
The intellectual functioning of a post-institutional group (Arthur D. Haggerty)	303
BOOKS	307
CRITICAL REVIEWS OF RECENT BOOKS	309
Humphrey, G. <i>Thinking</i> . (Reviewed by Walter B. Pillsbury)	309
BOOKS RECENTLY RECEIVED	313

CHILDREN'S CONCEPTIONS OF SOME BODILY FUNCTIONS*

Radcliffe College

MARIA H. NAGY

A. PROBLEM

After investigating children's conceptions of death (4) the author proposed to investigate their ideas relating to life. There are two main monographs on children's ideas of natural phenomena, by Piaget (6) and Oakes (5). Both dealt more with the conceptions of physical phenomena than biological; very little attention was given to the conceptions of the human body. The author therefore restricted her present investigation to the conceptions of human body functions. Since the human body performs a great number of functions the author limited the study to three: (a) the functioning of the brain, (b) breathing, (c) digestion.

This selection of topics was made on the basis of a preliminary investigation. When answering the instruction, "Tell me all that happens inside your body," children dealt mainly with these three functioning systems. It goes without saying that they were also interested in problems relating to birth. The author intends to publish her findings on birth conceptions in a separate study.

Since Schilder (7, 8) introduced the term "body image" into clinical psychology, there has been a growing interest in the body image. Although the body image seems to be a unitary phenomenon, historically the term implies a dual problem to be solved, first the question of the priority between the whole and the parts, and second the relationship between the intellectual and the emotional aspects of the body image.

The author is dealing with the conceptions of some functions of the body and not with the body as a static whole. The hope is that this analytic and dynamic approach will provide some new insight on the concept once the conceptions of the main bodily functions have been analyzed. This should occur to some extent in that the analysis is from part to whole, an approach differing from the direct analysis of the whole as such done by Schilder (7) and most of the clinical psychologists recently, e.g., Bender and Keeler (1). One instance of a different approach from these is Machover's (2). She has analyzed the different parts of the body in drawings, but without proposing

*Received in the Editorial Office on August 9, 1951.

any general conclusion on the body image. Furthermore, she deals mostly with the static aspect of the body, with the organs and not with their functions.

Murphy (3) thinks that the body image is not overwhelmingly important on the cognitive side, but on the affective as various parts of the body become associated with the self which is slowly differentiated from the outer world. Schilder's (7) point seems to be similar: "The study of the body image lies on the impressive side of our psychic life." Since the cognitive side of the body image has thus been taken into consideration less in the past than the affective side the author proposes to study this cognitive aspect.

Most of the previous studies of the body image were done with *clinical* subjects. The author used normal, non-clinical subjects for two reasons: (a) the attitude of the normal person toward his own body is also a basic determinant of his whole behavior and plays a part in his personality development, so that with age the individual organism is persuaded to act more and more like a unit; and (b) this investigation might lead beyond purely theoretically significant data and be helpful in the teaching of biology and hygiene. The methods and topics of biology and health education cannot be well selected if the assimilation of conceptions from the environment has not been checked.

B. PROCEDURE

1. *Interview With Drawing*

After an introductory talk to make contact with the child, during which he was offered candy, the experimenter (*E*) asked one of the following questions: (a) What happens inside your head? (b) What happens inside your body when you breathe? (c) What happens to the food that you eat? These questions were discussed on three different sessions in one week's time. The length of each session varied from 15 to 30 minutes. No child was forced to come or to stay. If there was any difficulty in starting the child speaking, he was asked to draw, and then to explain his drawing. At an appropriate moment it was suggested to each child that he make a drawing of the organ(s) in question. The interviews followed the general rules of interviewing, that is avoiding suggestion and allowing as much spontaneous response as possible. During the interviews *E* restricted herself largely to such comments as: *How do you mean?*, or *Tell me more about it*. One hundred sixty children age 4.0-11.11 served as subjects in Budapest, Hungary, in 1947-48.¹ Twenty children were taken for each year of chronologi-

¹Throughout the paper ages will be given in years and months (YM).

cal age. The distribution according to sex was: girls 13 per cent and boys 87 per cent. All the children were from the same school, which had kindergarten and elementary sections. The mental and socio-economic levels were estimated average by the school authorities for the total Budapest school population.

2. Essay Writing

Three essays were written, one each day for three days. The topics were the same, but their formulation differed slightly from the interview questions: (a) What happens inside our head? (b) How do we breathe? (c) How does the food travel from our mouth? The form to be filled in contained the following instructions: "Please will you answer a question? This is not school work and your writing and spelling will not be corrected. Do not ask anything and work on your own." Time allowed was 20 minutes for each question.

The age range covers the whole elementary school group except the entrance class, which was too young for any essay writing. Chronological age range: 8.0-10.11. The 270 children were equally divided within each year of the above age range. Children were from an average elementary school, Bristol, England, in 1949-50.

3. Tests 1-3

To establish Tests 1-3 they were presented to 100 children individually in the Belmont Day School, Belmont, Mass., and the Page School, Wellesley, Mass., in October, 1950. The main investigation was carried out in the Daniel Butler School, Belmont, Mass., in November, 1950, with 220 children. For the composition of the group see Table 1.

TABLE 1
DATA RELEVANT TO SUBJECTS OF TESTS 1-3

<i>N</i> = 220	Ranges	Overall <i>M</i>
1. Chronological age	5.2 - 11.3	—
2. Mental age	5.11- 12.2	—
3. <i>IQ</i>	106 -114	110.43
4. School achievement	2.5 - 2.55	2.38
5. Socioeconomic status of children	2.1 - 2.5	2.33
6. Occupation of the parents	3.0 - 3.85	3.48

Notes:

2 & 3. Stanford Binet Test.

4 & 5. Scored by the teachers on five-point scale where 1 is the highest value.

Here the class averages are given.

6. Sims' classification (10) on the above five-point scale.

7. Sex was not taken into consideration as it did not prove to be a significant variable.

Tests 1-3 compiled into one were also distributed to 300 parents at the same Belmont public school at a *PTA* meeting in January, 1951. Although no identification was required, only six were filled in and returned.

a. Test 1 consisted of an 8.5- x 5.5-inch booklet with seven pages. Children read the following instructions aloud with the teacher:

On the following pages you will find six circles. Draw six different things which are inside your head in these circles. Put only *one* thing in each circle. Do not ask anything. Work on your own. Your spelling and writing will not be corrected. Do not turn the page over unless you are told. (Time, 1 min. for each.)

Page 2 was as follows: Draw in one thing which is inside your head. Its name? What is it for? What is it made of?² (Time, 20 sec. each.)

Pages 3 to 7 were the same, and the following text closed Test 1: "Thank you very much for these drawings; it is hoped that you have enjoyed making them."

Two main objections might be raised against Test 1, first that the instruction speaks of circles whereas the schemes of the head were oval. This was done deliberately, because the children seemed to understand the word "circle" better than the words "figure" or "oval"; second that the contents of the head might be both physiological and psychological. If the purpose of the investigation was to find out which physiological items are known by children some allusion should have been made to this fact. But investigations *A* and *B* showed that the lengthening of instructions in this direction troubled the children.

Elementary school, Grades III-VI, did Test 1 in classes under the direction of the teachers. Since in this school the psychological service has been very extensive, neither the teachers nor the children needed any special preparation for Test 1. Kindergarten and Grades I-II volunteered for the test, and it was administered individually by *E*.

b. Test 2 contained very similar instructions to those in Test 1. The only difference was that the subjects of the drawings were given as follows: brain, nerves, bones, blood, sense-organs³ and everything that is in the head and not yet drawn. These items were taken from the essays. Test 2 was given to Grades III-VI at the end of the same morning school session as Test 1. Kindergarten and Grades I and II took Tests 1 and 2 separately, about a week apart.

²The diameter of the head scheme was 4 x 2.65 inches.

³"Sense-organs" were used since children confused the "senses" with the expression "to make sense."

C. CONCEPTIONS OF THE BRAIN

Up to age 7 the majority, 62.07 per cent, of the interviewed subjects and essay writers mentioned the brain as being inside the head.⁴

In the drawings *closed* figures appeared oftener than *open* ones. Of the closed figures, *round* figures outnumbered *angular* ones 6:1.

TABLE 2
DRAWINGS OF THE BRAIN

	Interview	Test 1	Test 2
1. Closed figures	71.96	69.15	72.5
2. Open figures	28.04	30.85	17.5
	<u>100.00</u>	<u>100.00</u>	<u>100.00</u>
	N/107	N/188	N/160

The brain was represented as made up of the following: bone, blood, skin, flesh, cells, and some miscellaneous items. "Cells" and "nerves" were given only by Grades V and VI, with frequencies under 5 per cent. The main constituent of the brain was considered to be bone, which shows a confusion between the brain itself and the skull.⁵

TABLE 3

	Test 1	Test 2
1. Bone	45.31	42.28
2. Blood	9.38	14.09
3. Skin	12.50	14.77
4. Flesh	4.69	10.07
5. Cells	4.69	4.70
6. Miscellaneous	23.44	14.00
	<u>100.00</u>	<u>100.00</u>

The brain was conceived as performing mainly intellectual activities. One might wonder how to explain the survival of this old conception in present everyday language. The explanation lies in the fact that the child is asked to use his brain when mental effort is required; such an appeal is never made in relation to his emotions or will.

Among intellectual activities the main one mentioned was thinking. In stereotype form: "The brain helps us to think." The occurrence of the item

⁴The alternative use of the words "brain" and "mind" in common Hungarian speech led children below eight years to think in terms of two distinct organs in the head with different functions. "Brain" was preferred by 75 per cent.

⁵Four of the six adult questionnaires returned represented the brain as scribbings covering the whole head. All four were college graduates.

"thinking" is four times as great as that of all other intellectual activities. It was only at 8 years of age that the interviewed children could talk about the activities of the brain without having any question about functioning. There were also general items on brain function, such as "working," "moving," etc. These imply lack of real knowledge as to specific brain functions; such answers decrease with increasing chronological age.

TABLE 4
BRAIN ACTIVITIES

	Interview	Essay	Test 1	Test 2
1. Intellectual activities	62.5	67.04	82.35	75.71
2. Non-intellectual activities	37.5	32.96	17.65	24.29
	100.00	100.00	100.00	100.00
	N/160	N/270	N/119	N/177

As the item "nerves" was only mentioned by 6.3 per cent among the contents of the head, in the essays; in Test 2 direct inquiry was made about nerves. It was found that the majority of the subjects represented the nerves graphically the same way they did the brain, namely, with both closed and open figures; again the closed roundish and angular figures predominated, 57.93 per cent from the 61.58 per cent given. The nerves were represented by threads covering all or part of the head by 38.42 per cent of the subjects. This represents the *specific* conception of nerves. The lack of distinction between brains and nerves might be the result of the children's global perception, but on the other hand, some constituents of the head were distinguished such as bone, blood, skin, and flesh. Of course these items are more in the sight of the children than the nerves, being outside the body.

TABLE 5
GRAPHIC REPRESENTATION OF THE NERVES

	Test 2, N of the items 164
1. Same as the brain	61.58
2. Different	38.42
	100.00

The functions of the nerves according to Test 2 results were: feeling, working, moving, thinking, and some vegetative functions. The main item is "feeling," referring mainly to such negative states as tenseness and nervousness. Also, 21.82 per cent of the items confused nerves with nervousness, and hence represented their functions as irritations of some kind. The fact that the main function of the nerves was considered to be different from that of

the brain should serve as evidence against the explanation by "globality," although the graphic representation of the brain and nerves was mainly identical.

TABLE 6
FUNCTIONS OF THE NERVES, TEST 2

1. Feeling	51.82
2. Others	48.18
	<u>100.00</u>
	N/110

The constituents of the nerves were enumerated in a list quite the same as for the brain. Bone was the main item mentioned, and the others were: blood, flesh, skin, cells, and some miscellaneous items with frequencies under 17 per cent.

TABLE 7
CONSTITUENTS OF THE NERVES, TEST 2

1. Bone	39.67
2. Blood	12.40
3. Flesh	11.57
4. Skin	16.53
5. Cells	3.30
6. Miscellaneous	16.53
	<u>100.00</u>

In short, the graphic representations of brain and nerves and also the list of their constituents were, according to Test 2, mainly similar. It becomes problematic how, this being the case, their functions can differ as they do, with the brain performing intellectual functions and the nerves emotional ones. The inconsistency here is traceable to the more general inconsistency in children's thinking, and to their slowly developing sense of contradiction.

D. CONCEPTIONS OF BREATHING

This section will not deal with children's ideas on the complete respiratory process, but only those on breathing as defined by the two complementary processes of inspiration and expiration. The intake of oxygen into the blood, etc., were not to be dealt with by the children.

In the essays on breathing the organs mentioned were mostly nose, mouth, and throat (61.03 per cent of the items). There was no mention here of what became of the intaken air after it reached the throat. This means that the air was considered to be circulating somewhere in the head region, more precisely between the nose and the mouth.

Also, 33.84 per cent of the total items of the essays mentioned a single organ in relation to breathing, chest, lungs, heart, or stomach. It seems that the children were inclined to localize the air in a certain part of the body without understanding the effect of the air on the whole body. There were mentioned here such items as a "bag," "box," "pipe," "tube," and "box" as the organs of respiration. On the basis of the entire essay this usually was intended clearly to mean the lungs. Blood, skin, and the body were other items mentioned in the essays in relation to breathing. They made up together only 5.13 per cent of the total items. This low frequency means that these children did not realize the rôle of the skin, blood, and body in breathing.

According to the interview results, children began speaking about breathing as a process at age six, but it is not until age group nine is reached that a predominant number (70 per cent) do so. This "process" did not include necessarily any interchange of oxygen, etc., but merely the taking in and expelling of air. Also, 53.7 per cent of the essay writers conceived of breathing as this static process, without any allusion to change, and 26.3 per cent did not give any data at all on the process, although the subject "How do we breathe" surely suggested it.

In investigating further the purpose of breathing by Test 3, it was found that although children affirmed the necessity of breathing they did not explain why it was necessary.

TABLE 8
THE RÔLE OF BREATHING, TEST 3

	No. 10	Questions No. 11	No. 15
1. Necessary for life	77.31	64.8	86.31
2. For taking air	18.87	35.2	8.33
3. Miscellaneous	3.77	—	5.36
	100.00	100.00	100.00
	N/159	N/125	N/168

TABLE 9
THE AIR INSIDE THE BODY, TEST 3

	No. 12	Questions No. 20
1. Moving around	74.04	63.21
2. Transformation of the air	21.26	25.79
3. Miscellaneous	4.72	11.00
	100.00	100.00
	N/127	N/106

In studying the ideas about breathing, other differently worded questions were included in Test 3. Again it was found that the majority of the children thought respiration to consist in a movement of air in and out or up and down, without any reference to change.

In Test 3 also two suggestive questions were given on breathing. The first suggested the sameness of the air breathed in and out, the second suggested the change in the process. According to Table 10, in this case the majority of the children accepted the change suggestion and answered accordingly. This implies that if the right ideas are suggested to children they are able to assimilate them, as here⁶ (Table 10).

As for the conceptions of the lungs themselves, it appears from both the essays and the interviews that children recognized only at the end of childhood that the organ specific to breathing is the lungs. In Test 3, after answering some questions on breathing children received some instruction on the lungs. These were given on a different page, separately from the questions on breathing. By this precaution the author hoped to avoid any suggestion connecting breathing with the lungs.

TABLE 10
THE CHANGES OF THE AIR IN BREATHING, TEST 3

	Questions	
	No. 16	No. 17
1. No change	41.58	6.45
2. Change	58.41	93.55
	100.00	100.00
	N/101	N/81

The question concerning the constituents of the lungs received answers of bone, skin, blood, flesh, cells, and several miscellaneous items. Again the same items were given as had been given for the brain, and bone was the main item, being mentioned 40.21 per cent overall. The frequencies for the other items ranged from 1 to 19.07 per cent. More general items, such as "cells," were mentioned only by the older children, about 3 per cent frequency. No allusion was made at all to the specific character of the lungs' constituents.

The question (19) inquiring about the function of the lungs was answered 74.23 per cent of the time in terms of breathing. The children knew the lungs' function, but as has been pointed out above, they referred much less

⁶Test 3, Question 17.

to the lungs when actually speaking about breathing. Perhaps it is not without interest to mention further that 7.98 per cent of the answers mentioned eating as a function of the lungs. In the earlier two investigations there were also some data reporting this belief. It might be explained by two suppositions: (*a*) some children considered the air as food for the body; and (*b*) the food has to pass through the chest and throat to arrive in the stomach and intestines, and therefore these children think the food must also pass through the lungs.

TABLE 11
GRAPHIC REPRESENTATION OF THE LUNGS

Test 3, Question 21, N/150			
A. Shape		B. Location	
1. Roundish	46.67	4. Above the chest	46.39
2. Veins	23.33	5. In the chest	30.46
3. Miscellaneous	30.00	6. Elsewhere	23.15
	100.00		100.00

The question of the lungs' activity (q. 20) was answered by the term "moving" in a significant proportion of cases, 63.21 per cent. These answers are in accord with the answers to the question (q. 12) which inquired about the breathing process. There are besides these general answers, "working," using, etc., further items attributing also talking and eating to the lungs.

Graphically the lungs were represented by roundish figures in about half the cases. Other representations were as veins, and about one-third of the items were so varied that no classification of them was possible. It may be noted that the children who drew lungs as veins did not draw a special organ in the center of the chest as the vein-center. Nearly half the children located the lungs above the shoulders, i.e., in the head or neck. Only one-third located them rightly in the chest. A few placed the lungs in the abdomen or outside the body scheme because they could not coördinate the lungs as a part of the body (Table 11).

E. DIGESTION

Regarding the passage of food from the mouth, three kinds of answers were given: (*a*) the food is going to the head and/or neck; (*b*) the food is going to the stomach; (*c*) the food goes beyond the stomach, because of being assimilated or evacuated. About half the children localized the eaten food in the stomach without any further explanation.

It is evident that most of the subjects knew only two parts of the digestive

tract, the mouth and the stomach. Neither the gullet nor the intestines was mentioned by a majority of the subjects⁷ (Table 12).

The interview was in certain respects a source of different materials than the essays, because of the different vocabulary. The everyday Hungarian language uses the terms "stomach" and "bowels" interchangeably for designating the organ of digestion. Some children used the two words alternatively, others did not distinguish between them. At about age 9 a significant number of the children recognized the stomach and the intestines as two coördinate, but different parts of the same tract. English and American children hardly mentioned "intestines," this term being far less common than "stomach" among them (Table 13).

TABLE 12
THE LOCATION OF THE FOOD INSIDE THE BODY

Level	Interview	Essay	Test 3
1. Head and neck	8.00	17.14	1.93
2. Stomach	51.7	65.08	43.75
3. Beyond the stomach	40.3	17.78	54.32
	100.00	100.00	100.00
	N/160	N/270	N/220

The stomach was represented by a roundish figure. This fact proves that children do not represent the digestive tract as a system composed of different organs. There was some tendency to locate the stomach higher in the body than is actually the case. This might be explained thus: In case the mouth and stomach represent the entire digestive tract the continuity of the mouth and stomach is better preserved by raising the stomach nearer the mouth (Table 14).

TABLE 13
RELATIONSHIP OF THE WORDS "STOMACH" AND "BOWELS," INTERVIEW

	Chron. age range	%
1. Alternative use*	4.0- 6.11	80.00
2. Identification	7.0- 8.11	47.50
3. Coördination	9.0-10.11	82.50
		N/160

*"Stomach" used 77 per cent and "bowels" 23 per cent.

"What is the stomach made of?" Neither the interviews nor the essays gave any relevant material on this problem without the direct question. The items given (q. 3, Test 3) were in the following percentages:

skin	bone	blood	flesh	cells	misc.
26.95	25.89	16.31	14.89	3.19	12.77

⁷Children in Great Britain (70 out of 160) used the expression "box," but from the essays it is clear they meant stomach.

As can be seen, the constituents given are the same as for the brain and lungs. No reference was made to the specificity of the stomach parts.

TABLE 14
GRAPHIC REPRESENTATION OF THE STOMACH, TEST 3

	Question 9	
I. Roundish figure in the region of		
1. Abdomen	8.37	
2. Chest	38.42	61.57
3. Shoulders	23.15	
4. Outside the body scheme	5.91	
	75.85	
II. Miscellaneous figures	14.15	
	100.00	
	N/203	

As far as the function of the stomach is concerned, it was found that the term "digestion" was not well known. The most frequently given answer is that the stomach is for keeping our food, or for eating. About 5 per cent of the children attributed breathing to the stomach. If this is compared with the answers relating to the lungs it may be concluded that there is no distinction made between certain inside organs, in this case between lungs and stomach.

In studying children's knowledge of the interior of the body Schilder and Wechsler (9) found that the whole inside of the human body is thought to be filled with food. In other words, since some inside organs are not clearly distinguished by the children they are inclined to locate food outside the digestive tract as well as within it (Table 15).

The idea that food fills up the entire body might also mean that the body is built up of foodstuffs. Consequently, Schilder and Wechsler's finding might be an expression of a notion about assimilation. The author herself has found children who were taught about assimilation, but who understood it

TABLE 15
"WHAT IS THE STOMACH FOR?" TEST 3

	Question 2
1. Eating and/or storing food	68.62
2. Digestion	14.89
3. Living	2.66
4. Breathing	4.79
5. Miscellaneous	9.04
	100.00
	N/188

to mean that the food in its original form spreads over the inside of the body. As one child stated it: "Only we do not see it because the food is inside our bones."

Answers on Test 3 (q. 5) about the *purpose* of eating showed that this idea is quite vague among children. They use expressions they have heard to explain it, but without any real comprehension of their meaning. To the direct question "Why do we eat?" the main answers were: to live, to keep healthy, to build our body, to grow, to prevent hunger, etc. These answers were quite scattered, having frequencies from 32.07 per cent to 5.91 per cent.

Since the word "digestion" was little used either in interview or essay answers, a special question was given to analyze its meaning. Slightly more children interpreted it to mean food storage than understood it as a transforming process.

TABLE 16
"WHAT DOES IT MEAN TO DIGEST OUR FOOD?"

	Interview	Essay	Test 3
1. Conservation of the food	53.29	52.19	58.70
2. Transformation of the food	37.50	47.81	41.30
	100.00	100.00	100.00
	N/138	N/228	N/152

For further analysis of the ideas on the digestive process consult Table 17. Conservation was said by the children to be the main part of digestion. They did not know that assimilation and dissimilation are complementary functions and cannot be separated from one another. They were unable either to unite the two-part functions into one system or explain their actual relationship.

In Test 3 there were three questions given as suggestions. Question 6 suggested the idea of assimilation; Question 7 suggested elimination. Question 8 suggested the correct notion by balancing the two-part functions.

TABLE 17
THE DIGESTION OF FOOD

	Inter- view	Essay	Test 3				
			No. 1	No. 4	No. 6	No. 7	No. 8
1. Conservation	63.45	87.26	91.35	96.38	95.33	12.39	40.00
2. Elimination	22.14	5.02	5.77	2.17	4.67	51.33	37.50
3. Both	32.86	11.72	2.88	1.45	—	36.28	22.50
	N/140	N/239	N/208	N/138	N/107	N/113	N/80

F. CONCLUSIONS

Here a brief attempt will be made to give some tentative explanations for the data presented above.

1. *Inadequate Environmental Stimulus*

The data had no positive correlation with the chronological and mental ages of the subjects. Since the investigation covers the chronological age range from 4 to 12, this fact needs to be explained. It seems that children do not develop their ideas on the subject according to the whole extent of their maturational capacity because the appropriate ideas were not given to them at all, or else not given in an appropriate way. Two considerations might support this statement: (a) 294 of the 300 parents did not return Tests 1-3, and the six questionnaires returned showed a high degree of ignorance on the material although four respondents were college graduates; (b) there are some book series on the subject for elementary school use, and they seem to provide the best basic knowledge of the matter. These, however, obviously do not achieve their complete aim, and this failure should be studied more closely. It is evident from the data of this investigation that the teaching of body functioning should be more intensive in the schools.⁸

2. *The Principle of Body Homogeneity*

These children had a single operating scheme with regard to the body, and they used it without regard to the specificity of organs. Brain, lungs, and stomach were all said to be of the same stuff: bone, blood, skin, flesh, blood, and miscellaneous items. Children thought the inside of the body to be constituted in the same way as the outer surface.

The concept of body homogeneity can be interpreted by the global perception idea, although this explanation seems inadequate. Strictly speaking, global perception would lead to the conception of the inside of the body as a group of organs not distinguished within the whole, or of organs built up of the same materials. Here the bodily organs and their constituents were differentiated, but thought to be identically constituted. In other words the organic constituents do not differ among the different organs, e.g., brain, lungs, and stomach.

⁸In the Belmont, Mass., public elementary school where the main study was done, these books were used: *Safe and Healthy Living Series*, Ginn and Co.; *New Health and Growth Series*, Macmillan Co.

3. *Graphic Representation of the Body Organs by Roundish Figures*

Children had no adequate knowledge of the shape of the relevant organs, but they wanted to please the *E*, so they drew something easy, not necessarily based on real knowledge. It would be worthwhile to investigate further why the roundish figure is a basic one for the representation of certain unknown things; also which kinds of things tend to be thus represented.

4. *The Body Functioning is Unilateral*

The transformations which take place within the body are little understood. In consequence, the two complementary processes, assimilation and dissimilation are not distinguished or understood. In those cases where one of these was mentioned, assimilation was mentioned most often.⁹

5. *Intellectualist Conception of the Brain*

What is the origin of the intellectualist conception still so prevalent in common language? "Use your brain" is continually repeated to stimulate thinking at home and at school. There is no reference to the brain in this manner when it is the child's will or emotions which are being encouraged. Why is the heart still considered to be the center of feeling? Which underlying human drive localizes certain human functions into different parts of the body? One might easily answer that this localization is due to the lack of differentiation between physiological and psychological factors (globality), but this answer would require a further explanation, viz., which psychological factors condition this globality? Can it be explained sufficiently by maturation or by the involved emotional factors?

6. *The Bodily Functioning is Static*

Each organ was attributed specific functions, but these were not described as a process of continuous change. Instead of describing the process itself children were inclined to define its rôle; the brain is for thinking, the lungs are for the air, the stomach is for the food, etc.

7. *The Functioning Systems are Oversimplified*

The brain was described as a single organ with one main function. Neither its relation to the nervous system nor to the body as a whole was taken into

⁹Ideas on elimination should be studied further because of their strong determinative influence on personality. The "anal" character was quite emphasized in psychoanalysis, but the developmental ideas of elimination were not then experimentally investigated.

consideration. Of the whole respiratory system only the lungs were mentioned, and they were said to have but one function. The relation of the lungs to the heart and circulatory system was also not mentioned. The digestive tract was simplified into one organ called the stomach.

8. *The Conception of the Self in Relation to the Body Image*

Since the approach to the body image was analytic and inductive, there is a greater need for an integrating principle. What is the relationship of the self to the separate bodily functions? (a) When the static aspect of the body is predominant, the self is related to the parts of the body through the intermediate experience of the whole body. (b) When the approach to the body is functional, the self seems to be related directly to the bodily function in question without relating to the wholeness of the body.

9. *Comparisons With Psychopathology and Anthropology*

Children's ideas of the bodily functions cannot be fully interpreted without the context of psychopathology and anthropology. Extensive investigations are needed in both fields to establish the basic components of the body image with the analytic and dynamic approach underlined here.

G. SUMMARY

Children's conceptions of some bodily functions were investigated with 650 children age 4.0 to 11.11. Three bodily functions were chosen: brain function, breathing, and digestion. Regarding each three problems had to be solved: (a) the graphic representation of the organ in question, (b) the constituent elements of the organ, and (c) the conception of the process itself. Some tentative explanations were also formulated, approaching the body image concept in a *dynamic* and *analytic* way.

REFERENCES

1. BENDER, L., & KEELER, W. R. The body image of schizophrenic children following electroshock therapy. Paper at the American Orthopsychiatric Assoc. Annual Conv., Detroit, 1951.
2. MACHOVER, K. *Personality Projection in the Drawing of the Human Figure*. Springfield, Ill.: Thomas, 1949. Pp. 181.
3. MURPHY, G. *Personality*. New York: Harper, 1947. Pp. 999.
4. NAGY, M. The child's theories concerning death. *J. Genet. Psychol.*, 1948, **73**, 3-27.
5. OAKES, M. E. *Children's Explanations of Natural Phenomena*. New York: Columbia Teachers' College, 1947. Pp. 151.
6. PIAGET, J. *The Child's Conception of the World*. London: Kegan, 1929. Pp. 397.

7. SCHILDER, P. *Das Körperschema*. Berlin: Springer, 1923. Pp. 92.
8. ———. *The Image and Appearance of the Human Body*. New York: International Universities Press, 1950. Pp. 353.
9. SCHILDER, P., & WECHSLER, D. What do children know about the interior of the body? *Intern. J. Psychoanal.*, 1935, **16**, 355-360.
10. SIMS, V. M. *Score-Card for Socio-Economic Status*. Bloomington, Ill.: Public School Publ. Co., 1927. Pp. 33.

12493 Cedar Road
Cleveland Heights, Ohio

CHILDREN'S BIRTH THEORIES*

Radcliffe College

MARIA H. NAGY

A. PROBLEM

Although this problem has been much discussed, it seems necessary to investigate it further: (a) with normal children age 4 to 11 from ordinary schools, (b) with more subjects, (c) with experimental methods.

The author seeks to discover and describe in children the development of birth "theories," or in other words, the children's conceptions of mammalian birth, on the basis of data from 390 children age 4 to 11.

Psychoanalysis first called attention to the importance of this topic in personality maturation as a whole and gave data from both free observation and psychoanalysis, mostly with clinical children. More recently Piaget (6) and Wallon (7) made major contributions on the topic by means of free observation of preschool children while studying children's theories on birth. Conn (2) was the first to employ the experimental method with clinical subjects, and he established some developmental stages in relation to the chronological age range within childhood.

Before going into detail the author must reflect on Barker's criticism (1) of her work on children's conceptions of death (5). To quote him: "Birth and death are among the most difficult intellectual problems that confront men, and they cause intense emotional reactions. It is very difficult to get at children's real conceptions of these matters." There are two kinds of emotional resistance that the *E* (Experimenter) meets at such times, namely social and topical. In reality these two kinds of resistance are quite inseparable, but the *E* has to attempt to relieve all social resistance to his personality. As to the topical resistance, or in other words, to the resistance against the topic "birth," the author does not think it possible to relieve it entirely in any short-time experimental situation. Furthermore she does not think it necessary to do so. The purpose of the present investigation is not to find out which birth "theories" children would have without any emotional resistance, but to analyze normal children's birth "theories" occurring in given experimental situations and formed in consequence of

*Received in the Editorial Office on August 22, 1951.

multiple factors, *including* emotional resistance. Children's *real* conceptions are those found in real life and influenced by all sorts of emotional resistance, not those that one would find in a hypothetical situation where there is no emotional resistance.

B. PROCEDURE

Two methods were used: (a) individual interview, (b) essay writing in school classes. Barker (1) criticizes the above methods when previously used by the author while investigating children's conceptions of death (5). He seems to connect the emotional resistance with the verbal techniques directly, although the creation of a free contact does not depend upon the verbiage of the method. Emotional resistance might be very strong despite some doll play techniques and might be relieved by an interview, of which the best example is the psychoanalytic interview.

1. *Interview* was used in two progressive kindergartens in Budapest, Hungary, in 1947-48 and in the Belmont Day School, Belmont, Mass., in 1950-51. In both cases the atmosphere was very free. Sixty Hungarian children age 4.0 to 6.11 years and 30 American white children age 5.4 to 10.1 years volunteered for the interview. They were equally distributed within the indicated chronological age ranges.¹ All had participated in other psychological experimentation before. The socio-economic standing of the parents in all these three schools was estimated to be average by the relevant school authorities.

After inquiring about the home and family of the child the following questions introduced the topic: (a) How did your cat or dog begin to live? (b) How did you begin to live?

The same questions related to members of his family. The interviewing followed the general rules for interviews, namely to avoid suggestion and to seek the maximum amount of free information. The *E* carried through the conversation mainly by such questions as "How do you mean?" or "What did you say?" Each interview lasted for about 15 minutes. If it happened that the child did not like talking with the *E* or got tired, he was sent back to his group.

2. *Essay writing* was administered by the teachers in their classrooms in three Elementary schools of Bristol, England, in 1949-50. Three hundred children age 8.0 to 10.11 years served as subjects. These schools were also considered to be average among the Bristol schools regarding the

¹Throughout ages are by years and months (YM).

socio-economic status of the parents. The written instructions on the test form to be filled in were:

We hear you enjoy answering questions, so please will you answer some for us? Write down everything you can about them. Do not ask anything and work on your own. We shall be pleased with all you write. This is not school work; your writing and spelling will not be corrected.

Half of the subjects in the largest of the three schools had been given the same two questions as the interviewed ones. Time was also 15 minutes for each question. The second group, 150 children from two small schools, had six questions. Two of these six were the same but differently worded and four further questions were added at the request of the principals, who feared parental criticism because of the topic. The six questions were:

- How did the world begin?
- How did a plant in the garden begin to live?
- How was my dog born?
- How is a bird born?
- How was I born?
- What makes me grow?

Time for Questions 3, 5, and 6 was 10 minutes each, and for the rest 5 minutes each.

C. RESULTS

Answers were divided into four classes and called "Theories" *A*, *B*, *C*, and *D* as follows:

- Theory *A*: There is no birth, as life has no beginning.
- Theory *B*: A mammal's life begins, but without the mother.
- Theory *C*: There is birth from the mother only.
- Theory *D*: Mammalian birth also implies the father.

The above four theories mark clearly some aspects of the mental and emotional development of the child from 4 to 11. As far as emotional maturation is concerned, the pleasure principle gradually gives place to the reality principle. Regarding intellectual maturation, the understanding of the time factor enables the child to conceive a beginning in time. Furthermore he will gradually be apt to take into consideration more and more interdependent factors in the birth process. First he explains the origin of the baby solely from the baby himself, then from the mother, and finally from the father also. So he exchanges his unilateral explanation for a bi- and tri-lateral relationship of the family.

Theories *A* and *B* were found only before eight years and Theory *D* ap-

peared only after eight years of age. The main theory was *C* in overall frequency. This means that children do not reach a basically adequate knowledge of mammalian origins in childhood. Whether this lack of knowledge is due solely to insufficient environmental stimuli, or is also determined by the maturation of the children cannot be answered definitely in this paper. It seems reasonable, however, to suppose that environmental stimuli do affect the child significantly only if he had previously reached a certain developmental level. From the present data it became clear that the majority of the subjects were able to understand the basic elements of the birth process only from about 8 years of chronological age.

Table 1 gives the frequency distribution of Theories *A-D* according to the three subsequent phases of investigation.

TABLE 1
BIRTH THEORIES
N = 426

Chronological age	4-6	8-10	5-10
A. Life without beginning	33.33	—	3.88
B. Birth without mother	11.67	—	0.97
C. Birth from mother only	55.00	77.95	72.82
D. Father included	—	22.05	22.33
	100.00	100.00	100.00
	<i>N</i> = 60	<i>N</i> = 263	<i>N</i> = 103

1. Theory *A*

The idea that there is no birth occurred in two variations: (*a*) some children stated that they had always been in the world or with their mothers, without any further explanation; (*b*) others gave a more detailed explanation about their different places of residence before coming to their present actual homes. Different residential places were mentioned, such as heaven, water, seashore, hospital, or different cities and countries. This implies that life was considered to last forever but that there were some stations distinguished in life according to past residential location. An example of the first of these variations follows:

Girl of 4.6 (S.F.): "The little cat has always been in the world."
 "And you? How did you get your mother?" "My mother has always had a daughter." "Who is she?" "Myself." "What happened when your mother was a little girl?" "She always had a daughter." "How?"
 No answer.

An example of the second type of answer under Theory *A* is the following:

Girl of 5.8 (H.B.): "I came from heaven." "How do you mean?" "I was in the Danube, in the water before." "How did you manage to go from the Danube to heaven?" "Jesus transported me there." "But how did you begin to live?" No answer.

The belief that life has no beginning and consequently no end reflects an attitude based entirely on the pleasure principle with exclusion of reality. Children of a certain age deny birth because they wish everlasting life; they are unable to accept any limitation in time. Theory *A* was also found in death conceptions of children by the author (5). Children on this level denied death, explaining it as a simple change of residence.

As to the time factor, the notion of a beginning presupposes the understanding of time as a continuous series, and the natural process in question as happening at a definite point in the time series, and so being limited and irreversible. Among previous investigators, Wallon (7) also called attention to the time factor in children's birth theories without referring it to the pleasure principle, although the denial of birth would remain unexplained without the context of the pleasure principle in the Freudian formulation (4).

2. Theory *B*

According to Theory *B* there is a beginning to mammalian life, but it is unconnected with the mother. Two variations were found: the beginning of life was described as (a) artificial, or (b) natural.

(a). The idea that the beginning of life is artificial means babies were fabricated, or manufactured. They can be made and perhaps bought in shops. God, man, or the stork were mentioned as makers. For example, one girl of 4.2 years (A.S.):

"How did your sister begin to live?" "The stork took a skin and stuck a head on. He made a point, a line, and a circle for the eye, and the baby was ready." "How could the stork get skin and bones?" "In the shop over the corner, towards the Danube." "It was very nice of you to have told me how a puppet is made; tell me how you were born." "The stork took me from a shop." "And your white cat?" "It is the same as for men." "How was it for men?" "The head, nose, eyes, legs and stomach were put on and it was ready."

The E related the birth theory to her, but she denied the truth of it. The artificialist theories were mainly found among Hungarian children and much less with Americans. The author does not find any explanation for that, as Dennis and Russell (3) had found artificialist theories among American children while investigating conceptions of other natural phe-

nomena. Piaget (6) and Wallon (7) also discovered artificialist conceptions of birth among Swiss and French children. Wallon explains them by a false generalization; children see that plants and some animals can be bought in shops and transfer these experiences to human babies. It seems, however, that this explanation needs completion, since children never saw babies bought in shops, at least not within Western culture. Furthermore, to buy plants and animals in shops does not necessarily include the idea of artificial manufacture. If one wants to find some real experience that suggests the artificialist birth conception to the child, one might think more of clay modeling. The child sees himself able to form human figures and, being unaware of the distinction between life and lifelessness, transfers his experience with clay figures to real human and animal babies.

(b). According to the second variation of Theory *B* birth is a natural phenomenon among mammals, but is unrelated to a mother. The origin of life is said to be a seed or an egg which just grows and becomes a baby. The egg or seed might be anywhere, on the seashore, on a small island, and so on. Children gave no further explanation as to the origin of the seed or egg. One boy of 5.6 (E.N.) explained as follows:

"And the calf?" "It comes from an egg." "What do you mean?" "It does not have horns yet to make a hole in the egg. The big cow makes one for it." "It is not true, you are telling me a story. The truth is that the calf comes from its mother's stomach." "No, it does not come from its mother's stomach." "Why not?" "It does not have horns yet to make a way out."

Analyzing the denial of the mother in mammals' birth from the standpoint of the pleasure-reality principle, the author's explanation is that the beginning of life is accepted in consequence of the growing reality sense, but the mother's rôle is denied, since to be born from the mother would mean dependence from her. The physical separation by birth and weaning aroused the child's aggression, and now he tries to find an emotional outlet in the denial of the mother. Furthermore, the child is not yet able to co-ordinate two factors in the same process, the mother and the child. So he operates in terms of a unilateral explanation originating the baby from the baby itself.

3. Theory *C*

According to Theory *C* the origin of babies is explained by reference to the mother, without including the father. Theory *C* is the main conception in all three groups used for the present investigation (vid. Table 1). Within Theory *C* two kinds of answers may be distinguished: (a) the origin of the

baby is explained solely by reference to the mother, and (b) some outside factors were also taken into consideration, but without relating to sexual intercourse. The ratio of (a) to (b) was 3:4.

(a). The first kind of answer explains the origin of a new mammal by factors immanent to the mother.

E.g., "This is just nature," "The baby was just born there in the mother's stomach," "An egg was formed inside the mother," "The mother gets it made," "The mother was born with the child inside," and so on.

In short, the mother's body is such that it produces the baby without any sexual intercourse.

(b). The second kind of answer is characterized by the intervention of an external person in the creation of a new baby. However, he is not the father, but God or the physician who puts something inside the mother—an egg, or a seed, or so on. For example, a girl of 4.11 years (M.F.) said:

"God puts the baby skillfully in so that the mother's stomach does not need to be cut. The baby grows and God takes it out nicely." "How do you mean?" "With his magic; he does it without cutting open her stomach." "What about the baby ape?" "The mother ape asks God for a magic almond; she is under magic; she eats the almond and the baby is born."

Analyzing Theory *C* from the standpoint of the pleasure and reality principle, the growth of the reality principle might be noted as children accept the mother, and consequently independence from the mother, although they are aware of never since possessing her so entirely as in the prenatal life and early babyhood. On the other hand they are not yet able emotionally to share the mother with the father, and therefore the father's rôle in birth is still denied. In Theory *C*, furthermore, they become able to relate the baby to the mother and give up the explanations originating the baby from the baby itself. This means they have exchanged their unilateral explanation for a bilateral one.²

The second group of interviewed children drew something that they wanted at the end of the interview: 83.78 per cent of the 37 drawings related to some containers such as a house, a car, an aeroplane, and so on. Taking the original Freudian interpretation this would mean these children symbolically referred to the mother. However, the meaning of drawings cannot be explained definitely without detailed analysis of each child.

²In an unpublished investigation by the author it was found that children of a certain age were unable to classify more than two words into one class. When further classification was suggested, they composed several dual units without any relationship among them.

Furthermore, any tridimensional object is to some extent a container, and this fact has to be taken into consideration when attributing symbolic meaning to any drawing.

4. *Theory D*

At the level represented by *Theory D* children accept full reality regarding the origin of babies. The father is included by these children in human conception. As Table 1 shows, *Theory D* did not appear before 8 years and did not become dominant in childhood. To show the extent of their knowledge at this level some of the best essays will be given verbatim:

Girl age 10.9 (P.H.): "I began as a very small egg in my mother's body when I was in the body of my mother she fed me by tubes. While my mother was feeding me I began to grow; first I grew slowly and looked rather like a fish and then after a long time I got to a normal size and I worked my way out through my mother's body until I got to an opening which I passed through. I could not live without a father because he has to give the mother some sperm out of his body."

A boy age 10.5 (P.F.): "Before a dog or cat could live the male and female animals had to combine to form the new being. It was done this way. Inside the female an egg was formed, then from the male came a little thing like a tadpole called sperm which swam into the mother's womb and penetrated the shell as soon as the sperm had done its work. The egg then began to take shape. Gradually the egg began to look like a small dog or cat which fed on food eaten by the mother. The formation of the baby became distinguishable, and at the end of a short period the baby was ready to enter the world."

Theory D children arrive fully at the level of the reality principle and accept the sharing of the mother with the father in birth. They also exchange their bilateral explanation for a trilateral one including the baby, mother, and father.

5. *Group Differences*

There are also two aspects which should be dealt with regarding the different schools used for the present investigation: the cultural aspect and the social one.

The children were from three different cultures, namely the American, British, and Hungarian. According to the data the cultural differences were not characteristic enough to treat the data according to them. Theories *A-D* did not distinguish the groups according to culture, but according to chronological age.

More differences were found between the schools from the social standpoint, specifically in the matter of sex education.

Hungarian children were officially taught less in sex matters than British and American children. They gave more fantasy material than real knowledge. However, this distinguished them only in the details of their answers and not in the essential features, since they gave the same Theories *A-D*. On the other hand the more extensive sex education of the parents in Great Britain and the U. S. produced more schematic answers. All the children were inclined to use the same words and expressions when speaking of mammalian origins without knowing very often the exact meaning of them. For example, an American boy age 9 told accurately the whole story of birth at the level of Theory *D*. For the designation of sexual intercourse he used the verb "to mate." When asked the meaning of "mate," he could not give any explanation. He had repeated what he had learned without having a concrete representation of it.

In Great Britain the three schools chosen for the investigation differed in their degree of prejudice toward sexual matters in the curriculum. One was very progressive, and sex matters were freely taught and discussed. The second was just the opposite, and the third was something between, i.e., an ordinary school. This classification according to sex prejudice was based upon the opinions of the elementary school authorities in Bristol, England. Table 2 shows the distribution of the answers according to these different

TABLE 2
BIRTH THEORIES OF 300 CHILDREN AGE 8-10

(A) <i>Subhumans</i>	Schools*				Total
	<i>M</i>	<i>H</i>	<i>L</i>		
1. From the mother	41.34	33.33	30.67		36.67
2. Father included	3.33	1.34	30.67		9.67
3. Miscellaneous**	23.33	65.33	10.67		30.67
4. No answer	32.00	—	28.00		23.00
	100.00	100.00	100.00		100.00
<hr/>					
(B) <i>Humans</i>	Schools*				Total
	<i>M</i>	<i>H</i>	<i>L</i>		
1. From the mother	44.00	12.00	26.67		31.67
2. Father included	2.67	2.67	30.67		9.67
3. Miscellaneous**	20.00	85.33	12.00		32.50
4. No answer	33.33	—	30.67		23.66
	100.00	100.00	100.00		100.00
	<i>N</i> = 150	<i>N</i> = 75	<i>N</i> = 75		<i>N</i> = 300

*According to degree of sex prejudice: *M* is average, *H* is high, *L* is low.

**Items dealing with the circumstances of birth, such as place, date, etc., not with the process itself.

schools. Schools low and high in sex prejudice differed clearly, and the school which was average in sex prejudice resembled School *H*(igh) more than School *L*(ow). Schools not only differed in the acceptance of the father's rôle in mammalian conception, but also in the frequency of the Miscellaneous and No answers. School *H* did not have the courage to refuse a direct answer, but gave the most Miscellaneous answers, dealing rather with the place and time of birth than with its process. It is also to be noted that the distribution of the items on subhumans and humans does not differ in the above indicated features.

D. SUMMARY

Three hundred ninety children from ordinary elementary schools were investigated by means of interview and essay writing regarding their conceptions of birth. Four "theories" were found, representing four levels of development: Theory *A*: There is no birth, as life is everlasting; Theory *B*: There is birth, but without the mother; Theory *C*: Birth is explained solely from the mother; Theory *D*: Father also included. Children from a less prejudiced environment in sex matters showed a clear advance in their birth conceptions. The four birth theories were tentatively explained by the gradual exchange of the pleasure principle for the reality principle in the Freudian formulation.

REFERENCES

1. BARKER, R. G. Child Psychology. In *Annual Review of Psychology*, 1951, **2**, 1-22.
2. CONN, J. H. Children's awareness of the origin of babies. *J. Child Psychiat.*, 1947, **1**, 140-176.
3. DENNIS, W., & RUSSELL, R. W. Piaget's questions applied to Zuni children. *Child Devel.*, 1940, **11**, 181-187.
4. FREUD, S. Beyond the Pleasure Principle. New York: Boni & Liveright, 1950. Pp. 90.
5. NAGY, M. The child's theories concerning death. *J. Genet. Psychol.*, 1948, **73**, 3-27.
6. PIAGET, J. La formation du symbole chez l'enfant. Neuchatel: Delachaux & Niestle, 1945. (Pp. 260-265.)
7. WALLON, H. Les origines de la pensee chez l'enfant. Paris: Presses Universitaires, 1947.

12493 Cedar Road
Cleveland Heights, Ohio

THE REPRESENTATION OF "GERMS" BY CHILDREN*

Radcliffe College

MARIA H. NAGY

A. PROBLEM

After investigating children's ideas on the origin of illness (1, 2) it seemed necessary to approach the problem from a different angle and make a separate study of their representation of germs. The word "germ" is used here, since this is the term used by children in Great Britain and the United States for designating any micro-organism causing infection.¹ According to these two former studies by the author (1, 2), the following conceptions of germs were found among children: (a) Children are inclined to attribute every illness to germs because they do not make the distinction between contagious and noncontagious illness. They do not refer at all to organic and functional diseases or dietary deficiencies. (b) They mention for the most part a single kind of germ; they are unaware of the fact that each disease has its special micro-organism. (c) The degree of infection and the resistance of the body to infection are left out of consideration. (d) The process of infection is considered to be an automatic one; if any infection gets into the human body it necessarily makes it ill. On the other hand, when medicine is administered into the body, healing takes place immediately.

In the present paper the author completes the above findings by investigating the following problems: (a) the graphic representations of germs, (b) the definition of germs, (c) the activity of germs, i.e., their entrance into, exit from and intermediate activity in the human body.

B. PROCEDURE

Four separate procedures were followed: (a) essay writing in school classes, (b) graphic representation of germs and interpretation of the drawings in groups, (c) group testing, (d) individual interview. Techniques *a-c* were used in one Bristol elementary school with children of 8 to 11 years in England in 1949-50. Technique *d* was applied to white American children of 5 to 7 years in a public kindergarten and elementary school in Belmont, Mass., in 1950-51. For further details regarding the subjects see Table 1.²

*Received in the Editorial Office on August 22, 1951.

¹In Hungary the term "bacillus" was current among children of 4 to 12; at the end of childhood the term "bacterion" was also used.

²The author has also aimed at collecting data from adults on the same topics, but

TABLE 1
CHARACTERISTICS OF THE TWO EXPERIMENTAL GROUPS

	Techniques 1-3	Technique 4
1. Number of subjects	340	30
2. Chronological age range	8.3-11.4	5.5-7.8
3. Sex distribution:		
a—Boys	44.4	40.00
b—Girls	55.6	60.00
4. IQ range*	89-131	82-140
5. Average IQ	106	114
6. Average school achievement**	2.5	2.87
7. Average socio-economic standing	3.23	2.27
8. Parents' occupation, average	3.00	2.9

*With children of 8-11 the average of three tests were used: (a) Richardson, C. A., *The Simplex Junior Intelligence Scale*, London: Harrap. (b) Schonell, F. J., and Adams, H. R., *The Essential Junior Intelligence Test*, Form B, Edinburgh: Oliver & Boyd. (c) Lawlet, D. N., *Moray House Intelligence Test 40*, University of London Press. With children of 5 to 7 the revised Binet was used.

**For 6, 7, and 8 five-point scales were used where 1 is the highest value. Six and 7 were scored by the school teachers and 8 by the author on the basis of Sims' occupational classification (3).

1. *Essay writing* was administered by the school teachers. The form to be filled in contained the following instructions:

We know you like answering questions. Would you answer one for us? This is not school work, and your writing and spelling will not be corrected. Do not ask anything and work on your own. We shall enjoy all that you write. The question: How do germs work in our body?

Time allowed was 20 minutes. During this time teachers did not allow any questions on the subject to be asked in order to avoid the possibility of suggestion.

2. *Graphic representation* of the germs was done a half year after the essay writing and with the same children. The test form contained the following instructions: "You have heard of germs in relation to some illnesses. Would you like to draw some germs?" After 10 minutes the test forms were turned over and the following instructions were read: "What is a germ? Please explain now what you have drawn." Time allowed for explanation was also 10 minutes. Teachers did not allow questions and placed the children far enough apart to avoid collaboration.

3. *Testing* was done the same day as the drawing of germs, in the first since her previous questionnaire on bodily functions was not filled in and returned, it was impossible for her to try anything more in the same area. However, she feels it very important for comparative purposes.

and last school periods of the morning. The test form contained the following instructions this time:

Please will you answer some questions for us. Write down all you can; do not ask anything and work on your own. The questions:

- (a). In what way do germs enter our bodies?
- (b). What do germs do in our bodies?
- (c). In what way do germs leave our bodies?

Ten minutes was given to answer each question.

4. *Interview* was used with children of 5 to 7 years, since they were too young for Techniques *a-b*. All the children volunteered, and *E* knew them all beforehand, as they had taken part in another investigation on bodily functions in three sessions. Time was about 15 minutes for each child, and the oral instructions were as follows: "I am going to ask some questions that you have not heard about. If you can answer them, fine. If you cannot it doesn't matter." The questions were: (a) Have you ever heard the word "germ"? What does it mean? (b) Would you like to draw some germs? (c) Through what ways do germs go into our body? (d) What are they doing inside? (e) When we are not ill any more, what happens to the germs?

C. RESULTS AND DISCUSSION

1. *Graphic Representation of Germs*

Nearly half of the children age 5 to 7 did not make any drawing of germs. The fact that the majority of the children in this age range were not able to make a picture of germs was confirmed earlier by an investigation of the author's with Hungarian children (2). In those instances in which something was drawn, it was either some abstract figure such as dots, some roundish or angular figures, or animals. Most of the drawings were abstract figures.

With children of 8 to 11 years, the drawings made may be analyzed as (a) wholes, or (b) in their elements.

a. *The synthetic approach.* If the drawings are considered as entities,

TABLE 2
DRAWINGS ON GERMS, TECH. 4

1. Abstract figure	43.33
2. Animals	10.00
3. No drawing	46.67
	<hr/>
	100.00
	N = 30

according to Table 3 three kinds of them may be distinguished: germs, animals, and scenes. This classification is based on the children's interpretation, since without interpretation no precise distinction would be possible between the pictures of the germs and animals. The class entitled "germs" does not need any further explanation. The class entitled "animals" does; it includes those items which are not called germs but animals, namely, flies, caterpillars, etc. The third class, "scenes," can best be illustrated by some verbatim examples given by the children.

One boy of 8.6 (G.P.) said: "I have drawn a bombed building with heaps of germs flying around, a rubbish dump with an old pair of shoes and some old tin cans with flies flying round it."

Another boy of 11.2 (H.B.)³ said: "Somebody is there in pajamas with chickenpox. The dots are germs which you get all over you."

It is not obvious in Table 3, but it may be noticed that the subjects under

TABLE 3
DISTRIBUTION OF DRAWINGS ACCORDING TO CONTENT (TECH. 2)

I. Germs	49.25
II. Animals*	29.43
III. Scene	21.32
	<hr/> 100.00
	N = 333

*Animals not identified with germs.

the conditions described choose figures of a single class to draw, namely germs or animals or scenes, without any interchange; Classes I-III are exclusive.

b. The analytic approach. Analyzing the drawings in terms of their elements, three groups of drawings were distinguishable: animals, abstract figures, and a mixture of both. The abstract figures are: points, circles, parallelograms, etc. Table 4 gives the frequency distribution according to school grades. Abstract figures seem to decrease with chronological age, giving way to animal figures. This fact is surprising, since the opposite might be expected, i.e., that concrete representations would give way to abstract ones with increasing age. This fact might perhaps be explained by the widening of the children's experience. Abstract figures were drawn when children had no concrete representation of germs, and animals were drawn when later identification of germs with animals was established.

Comparing further the drawings of germs and animals, it was found that "germs" is used much more as a general term than is "animals." To speak

³Throughout the paper ages will be given in years and months (YM).

TABLE 4
CLASSIFICATION OF DRAWINGS ACCORDING TO FORM

	Tech. 2			Tech. 4	
	III	IV	School grades V	Total	Kg. and I-II
1. Animal	31.09	44.83	69.38	47.14	18.75
2. Abstract figures	52.10	43.10	18.37	39.04	81.25
3. Both	3.36	11.21	5.10	6.61	—
4. Miscellaneous	13.44	0.86	7.14	7.21	—
	100.00	100.00	100.00	100.00	100.00
	N = 119	N = 116	N = 98	N = 333	N = 16

of "germs" generally indicates a probable ignorance of the fact that every illness has its own micro-organism (2: 11, 12). There were only seven cases at the Grade V level in which reference was made to the specificity of germs. There was a quite stereotyped form introducing the interpretation of the drawings: "I have drawn some germs." The use of the plural did not mean a variety of germs, but their multitude within one single species. This is not the case with animals. According to Table 5 the general term "animal" was used much less frequently.

TABLE 5
GENERALIZED AND REPEATED DRAWINGS,* TECH. 2

	Generalized	Repeated
1. Germs	85.19	91.40
2. Animals	14.81	8.60
	100.00	100.00
	N = 108	N = 93

*"Generalized" means to draw such non-specific things as "germ" and "animal."

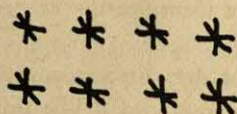
Table 5 shows also another fact which confirms the first one. In addition to the use of the general term referring to germs there was found a multiple graphic representation of the germs in identical form; children were inclined to draw "germs" in repetitive pictures. This was not the case with animals. In most cases children drew a single animal of a definite species, such as a fly, a worm, etc. The fact that "germ" was a more general term for the children would confirm the author's tentative explanation that for many children the word "germ" is meaningless (see Figure 1).

2. Definition of Germs

There were two kinds of material relevant to the definitions of germs: (a) the definition of the term "germ" in answering the question "What is a germ?", and (b) the interpretation of the drawings by the children.

(a). In analyzing the definitions of the concept "germ" the answers were classified under three headings by both age groups of the children. A germ was identified with some animals, or defined by its functions. The rest of the items under the heading "Miscellaneous" contain identifications of germs as dust, poison, etc. The main definition is in terms of function, for the most part stereotyped in the phrase, "Germs make us ill."

1.



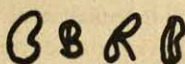
2.



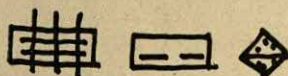
3.



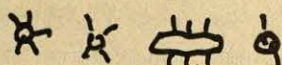
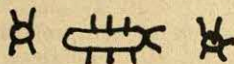
4.



5.



6.



7.

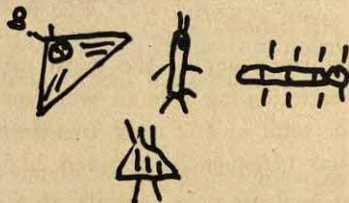
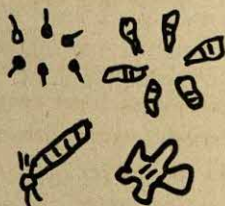


FIGURE 1
SAMPLES OF GRAPHIC REPRESENTATION

TABLE 6
DEFINITION OF GERM

	Tech. 2	Tech. 4
1. Identification with animals	33.65	10.00
2. Definition by function	62.02	60.00
3. Miscellaneous*	4.33	30.00
	100.00	100.00
	N = 208	N = 30

*Such items as "poison," "dust," "sickening stuff," "dots," and so on.

Two examples follow which show the identification of germs with animal species.

One boy of 10.9 (S.P.) said: "I have drawn some germs. The first one is a water spider. The second one is a mosquito. The third one is a kind of flea."

A girl of 9.4 (R.S.) said: "The fly is a very bad germ that carried a lot of disease. An ant is another dirty germ. A waterflea is another dirty thing."

It is clear in these cases that the germ was not conceived as a special kind of animal, but identified with some definite animal species. The animals which were believed to be germs were as follows, according to zoölogical classification. According to Table 7, Class II, Arthropoda, was the main

TABLE 7
ZOÖLOGICAL CLASSIFICATION OF ANIMALS IDENTIFIED WITH GERMS

	Tech. 2
I. Annelida	12.34
II. Arthropoda	74.89
III. Mollusca	3.68
IV. Vertebrates	8.23
V. Miscellaneous	0.86
	100.00
	N = 462

class. Among the Arthropods the *insects* were mentioned most often (289 out of 346 items).

If the frequency of the items rather than the zoölogical classification is used, the following animals were mentioned with over 5 per cent frequency from the total of 462 items in their rank order: fly, worm, flea, spider, ant, caterpillar (frequencies between 22.51 and 5.41 per cent).

After analyzing the identification of germs with animals, the definition of germs by function has to be analyzed next, on the basis of Class 2 of Table 6.

This is the main class referring to the definition of germs. Since the stereotype form "Germs make us ill" related to the different illnesses, these items gave a good indication of the children's knowledge of the range of diseases: 333 subjects gave 252 items on the various illnesses and enumerated 43 kinds of illness. The five main diseases were: chickenpox, measles, scarlet fever, colds, skin troubles. The frequencies on these ranged from 10.71 per cent to 6.75 per cent.

In Class 3 of Table 6 the germs were mainly defined as "things." One girl of 8.9 (Ch.B.) "Germs are very nasty things. They are things that look like little round dots." Other definitions refer to the germs as dust, dirt, poison, etc. One boy of 9.3 (J.K.) said: "I should imagine they are like little tiny pieces of dirt that we cannot see in the air. Sometimes they bring great illness to you."

(b). Analyzing the drawings on the basis of their interpretation, a different classification was reached because of the inconsistency of the children's statements; in this case between the name of the drawing and their interpretation of it.

On the basis of the interpretation of the animal drawings the relationship of germs to animals might be: (a) identification or likeness, (b) connection through illness, by such a statement as: "the germs and animals make us ill," (c) simple juxtaposition between germs and animals in cases where children mentioned both without mentioning any relationship between them, and (d) miscellaneous items in which the above three relationships were mingled. Examples of this last class are: One boy of 8.4 (T.G.), who said: "I think germs are little insects with legs and something like ants, only much smaller." Another boy of 8.4 (A.P.) said: "What I have drawn is a germ of a worm."

The main class is Class C, amounting to about half of the items. This fact shows well the confusion regarding the relation between germs and animals. Analyzing the origin of this confusion, it might be tentatively ex-

TABLE 8
EXPLANATION OF ANIMAL DRAWINGS WITH REGARD TO GERMS

	Tech. 2
a. Identification or likeness	16.06
b. Connection by illness	29.59
c. Juxtaposition	52.98
d. Miscellaneous	1.37
	<hr/> 100.00
	N = 436

plained by the spatial proximity of the two. Animals are often carriers of germs, as children learn from posters and other forms of health propaganda. Being at the same place they will be considered to be identical, or at least similar. This way of reasoning is also global, but the author prefers to formulate it more specifically and call it "the principle of spatial proximity."⁴

3. *Activity of Germs in the Body*

There were three questions on this topic; they concerned: (a) the entrance of germs into the body, (b) their work inside the body, (c) the exit of the germs from the body.

Table 9 shows the distribution of the answers. The problems of entrance and inside activity aroused slightly more answers than that concerning the exit in both age groups.

TABLE 9
TOPICS CHOSEN BY THE CHILDREN ON THE ACTIVITY OF GERMS

	Tech. 1	Tech. 3	Tech. 4
1. Entrance into the body	49.49	35.61	35.19
2. Inside activity	39.29	39.15	37.04
3. Exit from the body	11.22	25.24	27.78
	100.00	100.00	100.00
	N = 392	N = 1,862	N = 54

a. *Entrance of germs into the body.* Children mentioned only three ways by which germs may enter the body: mouth, nose, and skin. Since the mouth is both the organ of breathing and eating, these two functions must be separated in considering germ entrance. According to the 467 data of Technique c, 79.29 per cent of the answers referred to eating as the way of getting infected by germs. Only 21.71 per cent said that germs get into the mouth while breathing. Answers of children from 5 to 7 years were not

TABLE 10
ENTRANCE OF GERMS INTO THE BODY

	Tech. 1	Tech. 3	Tech. 4
1. Mouth	63.77	70.44	63.16
2. Nose	24.15	16.89	31.58
3. Skin	12.08	12.67	5.26
	100.00	100.00	100.00
	N = 207	N = 663	N = 30

⁴In an unpublished research of the author's it was found that children of about 6 or 7 classify all the objects which are at the same place into the same class. E.g., "These are all on the table."

computed here, since the total number of their answers was very small (Tech. *d*).

The idea that the main way of getting infected is via the mouth might be explained by the accentuated orality of the children and their cathexis of their mothers. According to their early experiences, the mouth is the main source of pleasure, as it connects the child directly with the mother. Later they still keep on believing that the main contact with the world is through the mouth. To formulate this belief we might call it "the theory of oral infection."

b. Inside activity of the germs. Answers fell into three main classes. Children from both age groups mentioned three kinds of inside activities: germs make us ill, germs damage the body, germs live in the body. It was necessary to separate Classes I and II, since for the children the "damage" done by the germs was not necessarily an illness.

TABLE 11

	Tech. 1	Tech. 3	Tech. 4
I. Making ill	31.34	51.58	30.00
II. Damaging the body	14.74	17.42	20.00
III. Living	53.92	31.00	50.00
	100.00	100.00	100.00
	N = 154	N = 729	N = 20

To show all the conceptions referring to the inside activity of germs the data from children age 8 to 10 were tabulated in detail (Table 12). In Class I there were two kinds of answers, one specifying the illness due to the germs and one speaking generally of the illness. In Class II there were also unspecific answers designating the germs as harmful creatures without any further explanation. As special damaging activities the poisoning and killing effects of germs were mentioned. Class III enumerates walking, eating, and breeding as life activities of germs, while at the same time being disadvantageous to the body. According to Class IV there was a quite large proportion of items not giving any direct answer to the question or not answering at all (Table 12).

c. The exit of germs from the body. Four ways were mentioned through which germs could leave the body: the mouth, nose, skin, and anus. These, with the exception of the anus, are the same ones mentioned as entrances. Children think the germs enter and leave the same ways. Here again the mouth is conceived to be the main opening for expelling germs. Therefore the "theory of oral infection" might be recalled as a tentative explanation here also (Table 13).

TABLE 12

	N	%
I. Making ill		
1. No specification	172	
2. Specification of the illness	204	
	376	36.86
II. Damaging		
3. No specification	59	
4. Poisoning	44	
5. Killing	24	
III. Living	127	12.45
6. Walking around	63	
7. Eating	130	
8. Breeding	33	
	226	22.16
IV.		
9. Miscellaneous	22	
10. No answer to the question	24	
11. No answer at all	245	
	291	28.53
		100.00
		N = 1,020

TABLE 13
WAYS BY WHICH GERMS LEAVE THE BODY

	Tech. 1	Tech. 3	Tech. 4
I. Mouth	48.72	34.89	46.67
II. Nose	28.20	30.00	40.00
III. Skin	7.69	14.47	—
IV. Anus	15.38	20.64	13.33
	100.00	100.00	100.00
	N = 39	N = 470	N = 15

Regarding the functions of the above organs for expelling germs, the following were mentioned: coughing and vomiting for the mouth, breathing and sneezing through the nose, and anal evacuation. The rest of the answers left unspecified the functions which expell germs.

Children referred to the above four openings of the body as exits for germs, but in most cases they did not try at all to explain the cause. This same situation was found by the author in studying the conditions of falling ill. Neither in the case of illness nor in that of healing were the conditioning factors considered. The kind and degree of infection and the heredi-

TABLE 14
FUNCTIONS OF THE BODY EXPELLING THE GERMS

	N	Tech. 3	%
I. Mouth			
1. Coughing	68		
2. Vomiting	63		
3. Others and unspecified	33		
	164		16.08
II. Nose			
4. Breathing	36		
5. Sneezing and/or blowing	92		
6. Others and unspecified	13		
	141		13.82
III. Skin	68		6.67
IV. Using lavatory	97		9.51
VI.			
10. Miscellaneous	50		
11. No answer to the question	79		
12. No answer at all	260		
	389		38.14
			100.00
			N = 1,020

tary and present status of the body were not mentioned.⁵ The children mentioned only one conception in connection with the healing process and the exit of germs from the body. This was the reference to the physician and/or medicine in getting well. The stereotype of this conception might be formulated thus: "The doctor gives us medicine and it pushes out the germs from our bodies." Some verbatim statements illustrating this conception follow: One boy of 9.10 (R.P.) said: "The medicine cures the body and washes the dirt away." Another boy, 8.7 (H.E.) said: "When the doctor gives you medicine, it clears the germs from inside you."

TABLE 15
FACTORS INTERVENING IN THE HEALING

	Tech. 1	Tech. 3	Tech. 4
1. Kind of infection	—	—	—
2. Degree of infection	—	—	—
3. Heredity	—	—	—
4. Resistance to infection	—	2.06	—
5. Medicine	12.06	15.78	16.67
	N = 275	N = 1,020	N = 30

The idea that the healing process takes place automatically because of the medicine agrees with the former findings of the author on the process of infection (2). The majority of children kept till the end of childhood the idea that if any germ reaches the body one necessarily falls ill without regard to the status of the germs or the condition of the body. This idea might be formulated as the principle of "automatic infection and healing." The human body is thought to react automatically; any germ or medicine is thought to provoke an instantaneous change.

D. SUMMARY

The graphic representation of "germs" the definition of germs and ideas on their activities inside the human body were studied by 370 school children age 5 to 10. Drawings and their interpretation, essay writing, test, and interviews were used.

The term "germ" was used by children for designating any micro-organism causing illness. From the fact that some animals are carriers of germs the majority of the children concluded that germs are animals, mainly insects. This did not mean that they are a special animal species, but that they were identified with definite animals such as flies, spiders, etc. If the germs were not identified as certain animals they were graphically represented in geometric figures such as points, circles, squares, etc.

Germs were said to enter the body through the mouth while eating, and also to leave the body through the mouth by most of the subjects. Their inside activity was said to be harmful because they lived to the detriment of the body.

On the basis of the above findings some tentative explanatory principles were formulated: (a) *identification by Spatial proximity* with regard to the identification of germs as animals, (b) *the oral theory of infection* due to the accentuated orality of the children, and (c) *the principle of automatic infection and healing*, in view of the fact that for the children the effect of germs in the body and also healing both seem to take place without any consideration of other conditions, either of germs or body.

For further studies on this topic the collection of data from normal adults and clinical subjects, both children and adults would be suggested. Furthermore, the comparison of the above conceptions with those found in different cultures might be fruitful. The body image cannot be interpreted fully without including the conceptions of illness and healing with all the underlying emotional factors.

REFERENCES

1. NAGY, M. H. Children's theories concerning the origin of diseases. In *Proceedings and Papers of the 12th International Congress of Psychology at Edinburgh*. London: Oliver & Boyd, 1950. (Pp. 96-97.)
2. ———. Children's ideas on the origin of illness. *Health Educ. J.*, 1951, 9, 6-12.
3. SIMS, V. M. Score-Card for Socio-Economic Status. Bloomington, Ill.: Public Schools' Publ., 1937. Pp. 33.

12493 Cedar Road

Cleveland Heights, Ohio

W.I.S.C. PATTERNS OF A SELECTIVE SAMPLE OF BILINGUAL SCHOOL CHILDREN*

Santa Barbara County Schools

GRACE T. ALTUS

A. THE PROBLEM

The question of alleged differences in intelligence between school children of Mexican descent and their non-Mexican peers has been a topic for many research studies (6). Such investigations have been generally consistent in indicating that the measured *IQ* of the former minority group is, on the average, some 10-15 points below that of the remainder of the school population. How much of this difference could be accounted for by environmental factors is not known, but it is apparent that the psychometric inferiority of the Mexican group cannot be considered real while such obvious handicaps as bilingualism are left uncontrolled. The present study is an attempt to compare the intelligence test patterning of a selective portion of the Mexican-descent bilingual minority to that of a unilingual (English-speaking) sample equated on the basis of age, sex, and performance *IQ*. Admittedly, there are many other factors such as socio-economic level, parental education, etc., which might optimally be controlled as well as possible deleterious influences operating on a measured *IQ*.

The selection of adequate measures of language and non-language measuring devices plainly presents one of the most serious obstacles in any attempts to investigate a problem of this sort. Mere test nomenclature may not identify valid differences between the two areas, as was indicated by Carlson and Henderson (3), who found that mean non-language scores on the California Test of Mental Maturity were not significantly different from mean language *IQ*'s for a group of Mexican-descent children living in Los Angeles. It is of course possible, though not likely in view of other studies, that these children did not suffer from a language handicap. On the other hand, it would appear that the non-language test of the *C.T.M.M.* is either more heavily weighted with language factors than would at first appear—in the oral instructions, left-right differentiation, etc.—or else that the test was in some other way inadequate for the purpose.

The Wechsler Intelligence Scale for Children (*W.I.S.C.*) was chosen for

*Received in the Editorial Office on September 1, 1951.

the present investigation for several reasons. It contains both a Verbal and a Performance *IQ* as standardized on the same population. Although there are oral directions for the Performance tests, the materials are largely self-explanatory and could be made clear even through pantomime when a child's knowledge of English is minimal. Further, although most studies of dull or feeble-minded children reveal tendencies for higher performance *IQ*'s, Seashore (5) indicated in the standardization data for the *W.I.S.C.* that on this test, such discrepancies were slight for groups at all levels of the intelligence continuum. Even in the truly feeble-minded group, the mean superiority of Performance *IQ* over Verbal was only two points. A similar two-point difference was found between rural and urban children, favoring the latter in the expected direction. Hence, one would expect only a negligible Verbal-Performance discrepancy even in the rural feeble-minded group. It would thus appear, at least in comparing groups, that the mean *W.I.S.C.* Performance *IQ* would serve as a reasonably good predictor of the mean Verbal *IQ* except in samples where such severely handicapping influences as bilingualism were in operation.

B. SUBJECTS AND METHOD

All children in the study were enrolled in the schools of Santa Barbara County, California, at some time during the academic years 1949-50 and 1950-51. Most of them had been referred to the Guidance Department of the school system for intelligence tests relative to preliminary screening for special classes for the mentally retarded. A dozen members of the bilingual sample, older and brighter than the former, were tested as part of another research surveying reading. It will be brought out, however, that their test patterns were similar to those of the total group.

A child was considered bilingual if his cumulative record indicated that Spanish was spoken exclusively, or that both Spanish and English were spoken, in the home. Obviously the extent to which the children spoke Spanish in social situations could not be determined adequately, and it is reasonable to suppose that there was considerable variation in the emphasis on one language to the exclusion of the other.

The majority of the unilingual group came from the same two sources. In fewer than 10 cases the primary reason for referral was maladjustment, with intelligence testing incidental to other action. Because of the method of selection, both samples are positively skewed and are in no way assumed to be representative of the total school populations from which they were derived.

The Wechsler Intelligence Scale for Children was given to all subjects by the writer. Table 1 indicates the variables on which the two groups were equated: age, sex ratio, and Performance *IQ*. Since the mean age was close to 11 years, one might suppose that the bilingual group, notorious for late

TABLE 1
DESCRIPTION OF GROUPS

	Unilinguals	Bilinguals	Diff.	<i>t</i>
Age in years	10.98	11.14	.16	.30
<i>SD</i>	3.05	2.64		
Sex ratio (M:F)	27:25	35:32		
<i>W.I.S.C.</i> Perf. <i>IQ</i>	86.43	84.01	2.42	.88
<i>SD</i>	14.40	15.75		

school entrance and retardation once enrolled, would be somewhat more retarded in terms of grade placement. However, the median school placement for both samples was grade four, indicating similar overageness in both groups. The same picture might not hold in communities where bilinguals constitute a different proportion of the population and policies of strictly academic advancement are enforced (8).

C. RESULTS AND DISCUSSION

Differences in total *IQ* and subtest means were computed for both samples, as may be seen in Table 2. While the difference in Performance *IQ* between the two groups had been shown to be slight and statistically insignificant,

TABLE 2
MEAN *IQ*'S AND SUBTEST SCALED SCORES ON THE *W.I.S.C.* AND THEIR DIFFERENCES FOR UNILINGUAL AND BILINGUAL GROUPS

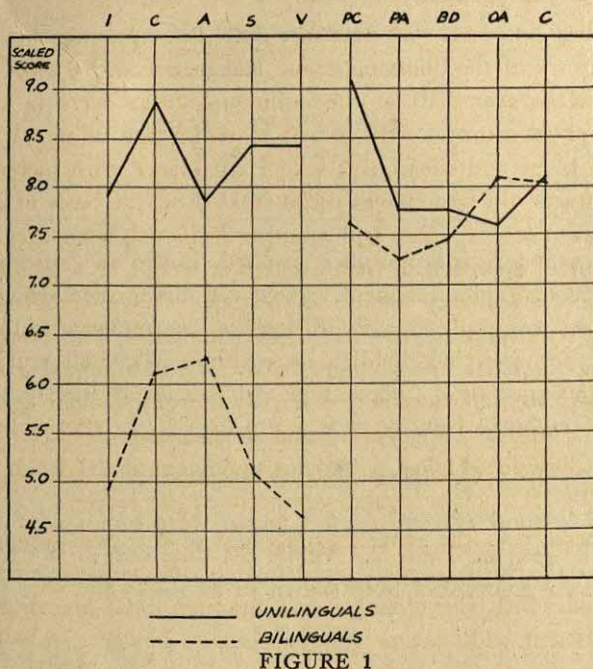
	Unilinguals		Bilinguals		Diff.†	<i>t</i> *
	Mean	<i>SD</i>	Mean	<i>SD</i>		
Verbal <i>IQ</i>	88.98	14.20	72.07	14.25	16.91	6.43*
Performance <i>IQ</i>	86.43	14.40	84.01	15.75	2.42	.88
Full Scale <i>IQ</i>	86.89	14.02	75.13	14.58	11.76	4.49*
Inf.	7.90	2.95	4.94	2.66	2.96	5.69*
Comp.	8.81	2.59	6.18	2.69	2.63	5.48*
Arith.	7.89	2.98	6.24	2.70	1.65	3.11*
Sim.	8.43	2.76	5.11	2.88	3.32	6.38*
Voc.	8.47	2.71	4.65	2.48	3.82	7.80*
P.C.	9.17	2.44	7.62	3.38	1.55	2.92*
P.A.	7.74	2.88	7.28	2.98	.46	.85
B.D.	7.74	2.72	7.46	2.43	.28	.58
O.A.	7.61	2.23	8.07	2.65	— .46	1.02
C.	8.10	3.06	8.05	2.95	.05	.09

*Differences significant at or beyond .01 level.

†Unilingual mean minus bilingual mean.

nificant, the average difference in Verbal *IQ*'s was nearly 17 points and highly significant, favoring the unilingual group as would be expected.

Figure 1 shows the *W.I.S.C.* subtest patterning for the two groups in graphic form. All verbal subtests show statistically significant differences



MEAN SUBTEST SCORES ON THE *W.I.S.C.* FOR UNILINGUAL AND BILINGUAL GROUPS

(.01 level of confidence or better) to exist between the two groups, with the most striking discrepancies on Vocabulary, Information, and Similarities, in that order. Although the bilingual group is thus within the psychometrically retarded range in verbal (English) skills, much of this retardation is apparently a linguistic one and does not reflect what the children could probably have done with training from birth in only one language—as evidenced more nearly, though of course very imperfectly, by the performance score. It is interesting to note that the English-speaking group follows Seashore's description in terms of a close correspondence between Verbal and Performance *IQ*'s.

The only subtest showing a significant difference between bilinguals and unilinguals on the Performance Scale is Picture Completion. This may be a result of the fact that a spoken answer is usually called for, and that even

though the child is credited with a plus if he indicates the answer clearly through pointing, many children in the earlier grades seemed reluctant to attempt an answer which could not be verbalized. A few first graders actually gave answers in English and Spanish interspersed, as though not fully aware of the distinction between the two languages.

One further note on verbal-performance discrepancies should be interjected. Only six of the bilingual group had higher verbal than performance *IQ*'s, whereas an even half of the unilingual group were in this category. In an attempt to determine the extent to which the bilingual discrepancies decreased with age and consequent school attendance with increased English-speaking contacts, the group was dichotomized on the basis of chronological age and patternings of older and younger bilinguals compared. Obviously the most fruitful approach to such a question would be a longitudinal study to see if school attendance narrowed the gap between verbal skills and the performance ones more representative of potential. The present technique of comparing the older and younger halves of the bilingual group is only a crude approach to this question, especially since the older group was also somewhat brighter on both Verbal and Performance *IQ*'s. (Their Verbal *IQ*'s were 9.39 points higher, difference significant at .01 level, and the Performance *IQ*'s 7.12 points higher, significant at .05 level.)

Figure 2 compares the subtest patternings of the two groups. It can be seen that on the verbal scales approximately the same relative position of subscales holds with the older group, although they are slightly brighter, as was mentioned. On the performance scales, however, it is interesting to note that on picture completion the two groups are widely separated, while rather close on three of the other four scales. Since this test is probably, as was noted, more heavily weighted with the verbal than any of the others, it may be that the older and younger groups are more nearly equated on performance skills than was originally thought. Hence, the higher verbal *IQ*'s of the older group may show that some tendency towards narrowing of the verbal-performance gap may occur with age and increased school attendance, although the relative difficulty of the subtests shows little change.

To test this idea further, performance-verbal discrepancies in raw numbers were plotted for both age groups of bilinguals. For the older bilinguals, the mean discrepancy was 11.18 points and for the younger, 13.0. The difference had an insignificant *t* of .71, but at least went in the expected direction. Had performance *IQ*'s been prorated after omitting picture completion, this tendency would have been somewhat more marked.

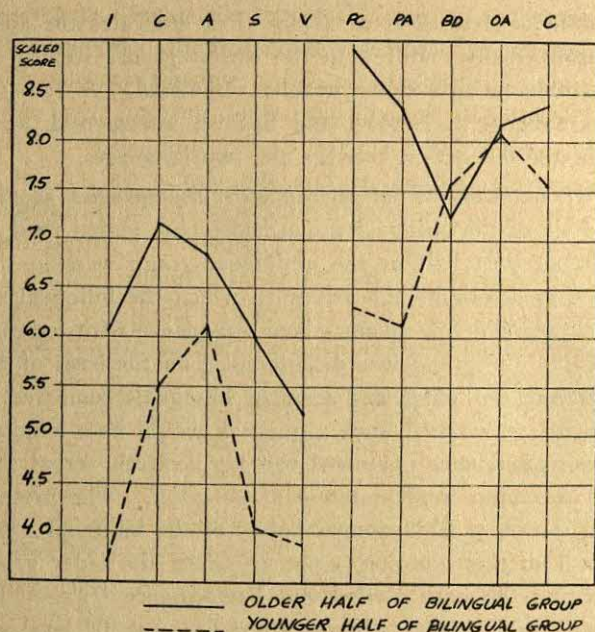


FIGURE 2
MEAN SUBTEST SCORES ON THE W.I.S.C. FOR THE BILINGUAL GROUP, DICHTOMIZED IN
TERMS OF AGE

There is evidence to suggest, however, that even if continuance in school might reduce the discrepancy between performance skills and English-speaking skills, termination of the formal educative process may in some instances halt or even reverse this process. A study by Altus (1) at a training station for illiterates dealt with an equally skewed distribution of adults. He reported that of all the Mexican-descent bilinguals present at the Training Center on a specified date, 18 per cent were literate in Spanish (while obviously illiterate in English) despite the fact that almost all had been born and educated in the United States, receiving their only formal instruction in English. Another 21 per cent were at least partially literate in Spanish. Because of the remarkable persistence of the linguistic preference herein revealed, it may well be that the Mexican-descent minority group suffers under a greater language handicap when using English than has been found for other minorities (4) and that the findings of this study would be unique to this cultural group.

Even within the bilingual Mexican-descent group, one would expect varia-

tions in intelligence test patterning for certain diagnostic entities. Altus and Clark (2) reported on the patterns of 31 juvenile delinquents belonging to this minority group. In this sample, 28 had higher Performance than Verbal *IQ*'s on the Wechsler-Bellevue, while *all* of the unilingual delinquents in the same study showed this characteristic. Subtest patterning showed similarities to be relatively high and arithmetic almost at the level of the vocabulary subtest. The performance patterning was also quite different from the bilingual sample herein described, with Digit Symbol sharply depressed and a peak on Object Assembly. The Mexican delinquent group actually followed a pattern highly similar to that of the English-speaking delinquent, neither of which resembled too closely the bilingual pattern herein described.

D. SUMMARY

Two groups of dull school children, one bilingual and of Mexican descent and the other unilingual and of non-Mexican descent, were equated on the variables of age, sex, and Performance *IQ* on the Wechsler Intelligence Scale for Children. Differences in *IQ* on the Verbal scale of this test averaged 17 points in favor of the unilingual group, a difference which was highly significant. A unique subtest patterning emerged for the bilingual group, which was dissimilar to the customary adult Wechsler pattern for the mentally retarded. It is felt that these results might be of value in the differential diagnosis of borderline cases of psychometric mental retardation within a bilingual Mexican-descent population, and might also offer some evidence as to the handicapping influences of bilingualism in this particular minority group. Since the sample involved was small and represented only one end of the intelligence continuum, it is apparent that no generalizations can be drawn in regard to all school children of Mexican descent, although it seems reasonable to suppose that verbal-performance discrepancies might exist at any level of intelligence. Further research, particularly of the longitudinal sort, is obviously needed to determine the magnitude of these differences and their accessibility to change as a result of continued schooling.

REFERENCES

1. ALTUS, W. D. The American Mexican: the survival of a culture. *J. Soc. Psych.*, 1949, **29**, 211-220.
2. ALTUS, W. D., & CLARK, J. H. Subtest variation on the Wechsler-Bellevue for two institutionalized behavior problem groups. *J. Consult. Psychol.*, 1949, **13**, 444-447.
3. CARLSON, H. B., & HENDERSON, N. The intelligence of American children of Mexican parentage. *J. Abn. & Soc. Psychol.*, 1950, **45**, 544-551.

4. KENDLER, T. S. Contributions of the Psychologist to constitutional law. *Amer. Psychol.*, 1950, **5**, 505-510.
5. SEASHORE, H. G. Differences between Verbal and Performance IQ's on the Wechsler Intelligence Scale for Children. *J. Consult. Psychol.*, 1951, **15**, 62-67.
6. TIREMAN, L. S. Teaching Spanish-Speaking Children. Albuquerque: Univ. New Mexico Press, 1948.
7. WECHSLER, D. The Measurement of Adult Intelligence. Baltimore: Williams & Wilkins, 1944.
8. YARBROUGH, C. L. Age-grade status of Texas children of Latin-American descent. *J. Educ. Res.*, 1946, **29**, 211-220.

Santa Barbara County Schools

P. O. Box 607

Goleta, California

INTELLIGENCE AND AGING*

Wisconsin Department of Public Welfare and University of Wisconsin

RAYMOND J. CORSINI AND KATHERINE K. FASSETT

A. THE PROBLEM

It is generally accepted that intelligence develops at a fairly steady rate from birth to about 15, at which time it either stops growing or else continues to increase at a very slight rate; at about 20 to 25 the absolute maximum of intellectual development appears to be reached. What happens to intelligence after this age is still open to question, although the consensus of modern opinion is that after this high point, intelligence begins a slow but steady decline.

It is the purpose of this study to review some of the more important large-scale investigations in this field, and to evaluate them in terms of our present standards of adequacy; also, to present and interpret some new data which we believe will suggest that intelligence does not decline in later maturity, or at least that the decline is at a much slower rate than usually believed.

B. THE MEASUREMENT OF INTELLIGENCE

It is axiomatic, in comparing two things, that the method of comparison should itself be valid. This requirement is not met in the case of intelligence tests, since their validity, especially beyond age 15, has not been conclusively demonstrated. An intelligence test may be a valid indicator for some classes of individuals and not for others; part of the problem which we shall soon encounter is the adequacy of a particular test at two levels of life. However, it is possible to come to some conclusions on the question of comparative intelligence by the use of tests, if proper cautions are used in interpretation.

C. OPINIONS AND CASES

Scientific truth can sometimes be forecast through popular opinion. With reference to intelligence and aging we find contradictory attitudes: on the one hand the aged are wise, and on the other hand the old become foolish and childish. Important advisory positions, be they in savage tribes or on the Supreme Court, are often filled by old men. But Shakespeare, in saying

*Accepted for publication by Harold E. Jones of the Editorial Board, and received in the Editorial Office on September 2, 1951.

of the final act of life, "sans everything," reflects the old saw, "once a man, twice a child."

Case studies also confuse us. While we do find elderly people who have reverted to childishness, we also do have examples of people of advanced age who maintain full intellectual vigor. Restricting ourselves to our contemporaries we need think only of John Dewey, Benedetto Croce, Oliver Wendell Holmes, Bernard Baruch, George Bernard Shaw, Winston Churchill, and Albert Einstein for evidence.

We must conclude, if we are searching for nomothetic law, that the two approaches of popular opinion and individual cases are inadequate, since evidence exists for both sides.

D. OBJECTIVE STUDIES

The first large-scale systematic investigation of intelligence and aging was done by Sir Francis Galton in 1876-77. Galton tested some 10,000 subjects in the age range of five to 80. But Galton's psychophysical measures are today in bad repute as indicators of intelligence. However, Charles Spearman, Galton's successor, commenting on the findings of Galton as well as the findings of Yerkes (*vide infra*), stated: "'*g*' general intelligence increases from birth—at first rapidly, then more and more slowly,—until somewhere not later than 15-16 years (and perhaps much earlier) its growth definitely ceases. Thereafter it normally retains this maximum level unaltered right up to the end of life (or, at least, to the onset of senility)" (10, p. 375). Spearman's conclusion is almost unique in the literature and is in direct contrast to Wechsler's statement that, "Every human capacity after attaining a maximum begins an immediate decline" (13, p. 55) and even more specifically: "Contrary to common belief, many of our intellectual abilities show greater impairment with age than do our physical ones" (13, p. 56).

Yerkes (15) reported on the scores of 15,385 white World War I officers. He showed that Army Alpha scores decreased with age. Miles and Miles (5) in 1932, using a shortened version of the Otis Test, found on a population of 823 subjects that scores began to decline at about age 20. Jones and Conrad (3) used the Army Alpha Test on 1,191 subjects in 19 New England villages and concluded that intelligence test scores declined in later years. These latter investigators attempted to control all possible non-intellectual factors involved. Wechsler (13) tested some 1,000 adults in obtaining norms for his Bellevue Test and concluded that his curve of intellectual ability was "very similar to the one that we derived from the

studies of Miles and those of Conrad and Jones" (13, p. 29).

It is apparent that these studies agree on one fact: intelligence test scores decline with age. But, as we have seen, these facts were interpreted differently by Spearman and by Wechsler.

E. EVALUATION

In evaluating this type of investigation two factors should be considered: the sampling of the population and the nature of the measuring instrument. If the populations at the various levels of age are not representative of the age population as a whole, or if the instrument used is inappropriate at any age, then the conclusions arrived at will not necessarily be valid.

In the Yerkes study, World War II army officers were used. The question arises, agreeing that as a class army officers are superior intellectually to the population as a whole, whether older officers are more or less superior to their age segment of the population than younger officers are to theirs. We have no way of determining whether unspecified selective factors exist for older men in comparison to younger men in the military, but the probability exists that some as yet unknown factors may be operating. Yerkes (15) comments that these results do not necessarily mean a fundamental dependence of intelligence on age but may reflect the selective process at work in separating the army from the general population, and Spearman, in commenting on this study, says, "The continual lowering of scores cannot be said, on the basis of present information, to point to a decrease of intelligence with age" (10, p. 372).

The Miles study and also the Jones and Conrad study may be criticized with respect to representativeness of population. Although large samples were obtained, we cannot be certain that they were not biased at the lower ages due to an excess of school-going youths, who tend to be brighter and more familiar with pencil-and-paper tests than non-school-going youths, and biased at the upper age levels because of an excess of retired, invalid, or financially depressed individuals. The classic study of Weisenburg, Roe, and McBride (14) points out unequivocally that older people resist voluntary psychological testing even under optimal conditions, so that any large scale voluntary sample must be regarded with some suspicion in terms of representative sampling.

A sample that is high at one end and low at the other may form a statistically normal sample when considered as a whole, but within such a population, valid comparisons between age groups cannot be made. It seems possible that something of this sort has happened within Wechsler's standardiza-

tion group: that he has a disproportionate number of students at the lower age levels, who tend to be superior to the national population, and that he has a disproportionate number of intellectually inferior individuals at the older levels. Wechsler's population, although obtained from 10 places (13, p. 112), has the greatest number selected from Federal Work Projects. On the assumption that these workers were in the older age brackets, we may feel, if relief workers tend to be inferior intellectually to their age peers, that the sample is biased. Further, Wechsler obtained a large sample of his older cases from institutions for the aged. This certainly does not constitute an adequate sampling of older people, since older individuals in institutions are undoubtedly more senile and less capable of adjustment than their age peers who remain outside of institutions. Bensburg and Sloan (1, p. 362) state: "The much higher *IQ*'s found on the Wechsler for older adults appear to be an artifact which may have been due to the inadequate sampling in his standardizing populations. . . . The present study having controlled this factor appears to cast considerable doubt on the validity of Wechsler's sample." In their review of research with the Wechsler-Bellevue, Rabin and Guertin state, however (7, p. 215): "It is rather obvious from the literature that the *W-B* is a well accepted measure of intelligence. Furthermore the standardization is equally well accepted. . . ." "The psychometrist's faith seems to be well placed in Wechsler's standardization and the test in general" (7, p. 216). "McNemar's enthusiasm for the use of Wechsler's standardization sample is understandable since it guarantees a proper range of talent . . ." (7, p. 218).

Let us now examine the tests used in these four studies. Except for Wechsler, intelligence tests used were of the group variety. Yerkes and Jones and Conrad used the Army Alpha while the Miles' used a special revision of one of the Otis tests. Conceding that group intelligence tests may have high positive correlations with individual tests, and that, while they may not be useful in ideographic clinical situations, they may be appropriate for the evaluation of group differences, the question arises whether group verbal tests are appropriate for between-age comparisons.

It is possible for any test to correlate perfectly with a valid criterion at a number of age levels but not to correlate highly with the group as a whole. This would be the case if, for the various age levels, certain non-intellectual factors which correlated with age but not with intelligence were to exist which would affect the test scores. This situation would not permit accurate between-age comparisons since the measurement would be contaminated.

To give a specific and pertinent example: let us assume that the Army

Alpha correlates 1.00 with a true criterion of intelligence at ages 15, 30, and 50, each taken separately. Let us further assume that the decline in visual acuity is positively correlated with age but not with intelligence. Then acuity becomes an element for the estimation of intelligence longitudinally in a test requiring visual acuity while it is not a pertinent factor at any single age level, since visual capacity tends to be equalized among individuals at any particular age level.

We believe that this is exactly what occurs in between-age research when group tests are used. There are a large number of non-intellectual factors which can depress scores to a greater degree among older people than young people. Among these are: recent experience with group tests, visual acuity, auditory acuity, manipulative dexterity, attitudes towards tests, etc. Consequently, unless we assume that intelligence is what intelligence tests measure, we must demand that a test used for between-age research be pure in terms of differential non-intellectual measurement—or to put this more simply, that a test should measure intelligence and nothing else which can be a function of age.

The Wechsler-Bellevue is an individual test, and it has become the instrument of choice in adult measurement. Most of the criticisms directed to group tests are not applicable to this test. Raw scores of the Bellevue subtests are convertible into equated scores. It also has some very important advantage in between-age research in that some subtests are verbal, some performance; some are timed and some untimed. Also, it provides a considerable variety of situations. While it may have serious faults for other purposes, for the purpose at hand, comparative age-group evaluation, it is an ideal instrument.

Summarizing the evaluation of the four large-scale investigations of intelligence and aging, we discover that none of them is entirely free of basic criticisms with respect to sampling and that three of them may be criticized with respect to the instrument used. At this point we feel safe in not accepting the facts of these investigations as necessarily proving that intelligence does decline with greater age.

F. AGE AND INTELLIGENCE—A HYPOTHESIS

We present as our hypothesis that intelligence does not decline significantly from early to late maturity, and that the decline noted by other investigators is mainly a function of two conditions: poor sampling, and the loading of non-intellectual factors in the tests. We believe that even if good sampling

in between-age studies is attained, decline can be hypothecated dependent on the degree of non-intellectual material in the criterion test.

We regard intelligence as a within-the-skin entity which manifests itself by overt behavior. To the extent that intelligent behavior is mediated through physiological—emotional—cultural vehicles, intelligence is measured with lesser degrees of accuracy. Consequently, we believe the amount of non-intellectual mediation governs the amount of intelligence test score decline at older ages.

Becoming more specific: with regard to the Bellevue Test, we have 11 subtests. We would hypothesize for older people that performance tests would drop faster than verbal tests, that timed tests would drop faster than untimed tests, that tests calling for speed and manipulation would drop most, and that tests calling for neither would drop least. We would guess that Vocabulary would be an example of a subtest which would be relatively unaffected by emotional-physiological factors while Digit Symbol Substitution would be greatly affected.

G. PENAL INVESTIGATIONS

We have criticized several population samples as to representatives at various age levels. We feel that Weisenberg, Roe, and McBride's findings (14) are significant in reference to lack of coöperation in adults, and that only forced samples will be representative. Prisons are perhaps the only place in our society where forced samplings of adult groups are possible.

How good a sample a prison population will be of the society at large cannot be known with any definitiveness. Sutherland's summary of mental testing in prisons (11) indicates that the trend of examinations has been to prove prisoners intellectually equal to the population in general. Although some of the early minor studies indicated that the prison population was not equal to the total population, three modern large-scale investigations, those of Murchison (6), Tulchin (12), and Schmidt (9), indicate that the prisoner is intellectually equal to the average of the social class out of which he came. Rouke (8, p. 454), in a summary of penological research, concludes: "The weight of evidence indicates that the distribution of prisoner and non-prisoner populations intellectually are almost completely superimposed." Dressler (2, p. 19) remarks: "A better understanding of scores of intelligence tests has brought the realization that criminals, taken in large numbers, will score about the same as the rest of the population if such factors as race, education, nationality, and location are taken into account."

It would seem that some authorities believe that prisoners as a whole

reflect a fair sampling of our total society in terms of intelligence. The questions we are specifically concerned with are these: do prisoners represent their age peers, or are the young prisoners brighter than the total young; and are the old prisoners duller than the total old, as we suspect of Wechsler's group, or is our sample biased in the other direction? Lacking evidence, we cannot decide, but we can see no especial bias that would affect age classes of prisoners in terms of intellect. In any event, a large scale penal population should point out whether subtest differences, as we have hypothesized, will show up with older people.

H. REPORTING RESULTS

Age-intelligence studies have reported results in two usual ways: correlations and graphs. The first has the disadvantage of the *N*'s usually tapering off at the upper age levels, so that the distribution is skewed, being overweighted at the lower age levels, and the tail we are chiefly interested in failing to get its full weight. Graphs are much better, but they are cumbersome and costly. In this report we will make use of a new statistic suggested to us by Kuhlen (4): the 15/50 ratio. This will tell us how well 50-year-old people do in comparison to 15-year-old people, defining them as the five-year periods beginning at 15 and 50. The decimal fraction 1.00 indicates parity, ratios below 1.00 indicate older people are inferior to younger people, and ratios above 1.00 indicate older people are superior to younger people. This simple statistic gives a quick understanding of trends, and, when the curve is regular, should be adequate. Since the Bellevue raw scores are converted into equated scores, this procedure is defensible. If the decimal is removed, the numbers may be regarded as per cents. We have taken these two ages because the younger is close to, if not right at, the maximum of mental ability, and because one can usually get large enough samples at both of these ages.

I. METHOD

The subjects for this report were incoming prisoners tested in the Guidance Center of the California State Prison at San Quentin from 1944-1948. Our original pool of scores was over 4,000 cases. The Wechsler-Bellevue Test, Form I, had been routinely administered, except for Vocabulary, to each inmate considered capable of understanding the English language, and over 99 per cent of entering prisoners were so examined. Our method of selecting cases was as follows: we took the first 100 cases in the age range 15-19, then started again with the first case in the age range 20-24, until we had 100 cases in that group. In the same manner we obtained 100 cases

for each 5-year period of age through 55-59. From age 60 on, there were fewer than 100 cases per 5-year period, but the magnitude of our sample in the later years is considerably greater than that reported for Wechsler's standardization population.

We have obtained means and standard deviations for each of the 10 subtests used in this study, for each of the 12 five-year age periods reported.¹ We have graphed our findings, as well as those reported by Wechsler, to permit comparisons. Finally, we have computed a number of 15/50 ratios.

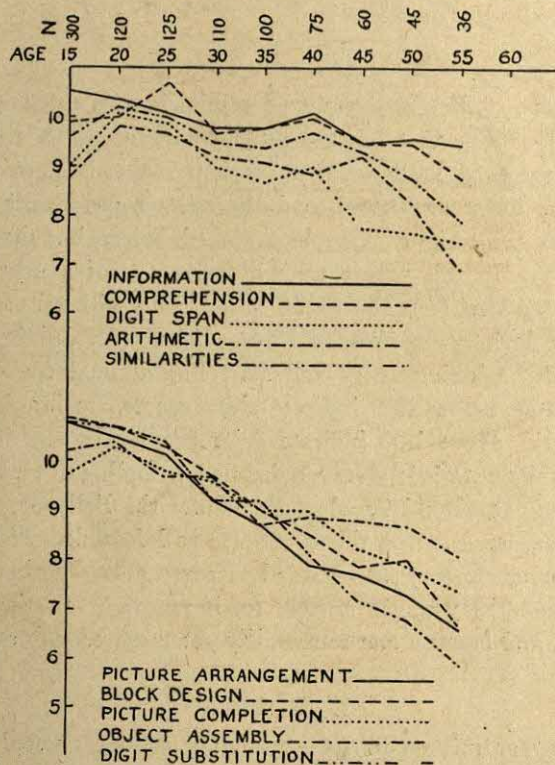


FIGURE 1

WECHSLER-BELLEVUE TEST, WECHSLER DATA: SUBTEST SCORES IN RELATION TO AGE

J. RESULTS

Figure 1, which has been graphed from Wechsler's own data, shows the general curves of decline for 10 subtests on the Bellevue Scale. We note several trends which have bearing on our hypothesis.

¹To reduce printing costs, Tables 1, 2, 3, 4 giving means and standard deviations

It can be seen that the performance material tends to fall at a faster rate than the verbal material. Also, it would seem that the performance tests are declining in a more consistent manner, while the verbal tests tend to spread. Digit Symbol Substitution drops the fastest, while Information holds best. Our hypothesis appear to be substantiated by Wechsler's own data.

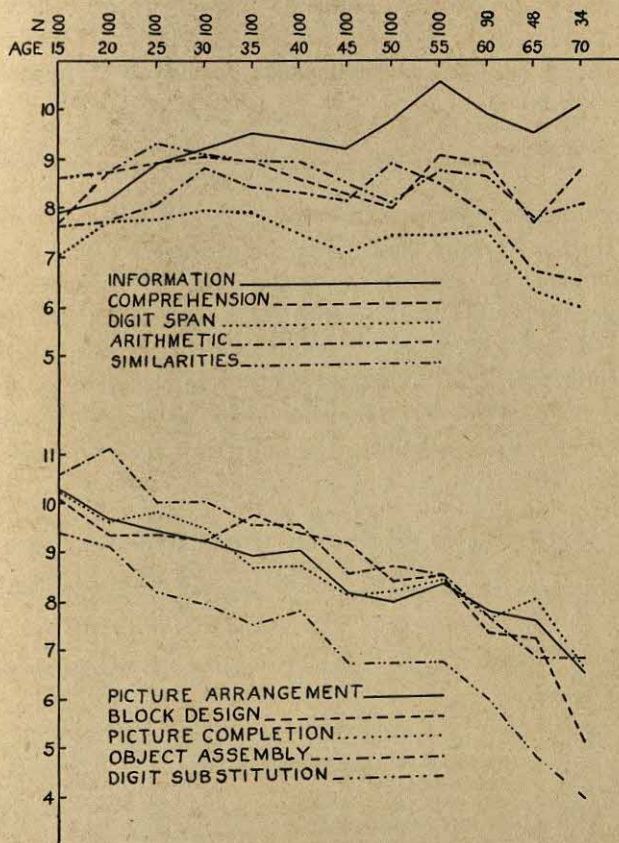


FIGURE 2

WECHSLER-BELLEVUE TEST, SAN QUENTIN DATA: SUBTEST SCORES IN RELATION TO AGE

The San Quentin data in Figure 2 shows the same trends: performance

of Wechsler's standardizing population as reported in his Tables 39 and 40, but reworked to conform to multiples of five-year ranges, for verbal and performance subtests; and for the San Quentin data, have been deposited with the American Documentation Institute. They may be ordered from the American Documentation Institute, 1719 N. Street, N. W., Washington 6, D. C., remitting \$0.50 for microfilm (images 1 inch high on standard 35 mm. motion picture film) or \$0.60 for photocopies (6 x 8 inches) readable without optical aid. The authors will supply the document number to any who request.

dropping faster than verbal, Digit Substitution dropping fastest of all, and Information least of all. But a very remarkable difference appears in these graphs: while the two graphs for Performance are about the same for both populations, the graphs for Verbal are quite different. Table 1 summarizes this difference.

TABLE 1
FIFTEEN/FIFTY RATIOS FOR THE WECHSLER-BELLEVUE

	Wechsler's data	San Quentin data
Verbal	.92	1.09
Performance	.74	.78
Global	.83	.92

On the performance subtests for Wechsler, the 15/50 ratio is .74, while it is .78 for the San Quentin data. This means that the older men in both populations do about three-quarters as well on the performance material as do the younger men. The 15/50 ratios for Verbal tell a different story: .92 for Wechsler and 1.09 for San Quentin. The nine per cent increase for the San Quentin data is in unexpected contrast to the eight per cent drop in Wechsler's data. While the difference in the performance ratios is but four points for the two populations, it is more than four times as large for the two verbal ratios.

TABLE 2
FIFTEEN/FIFTY RATIOS FOR THE WECHSLER-BELLEVUE SUB TESTS

	Wechsler's data	San Quentin data
I	.91	1.25
C	.96	1.05
A	.94	1.17
D	.86	1.07
S	.92	.97
PA	.67	.77
PC	.82	.84
BD	.74	.79
OA	.85	.81
DS	.63	.69

Table 2 lists subtest 15/50 ratios for the two populations: 15/50 ratios of less than .89 or more than 1.11 are significantly different from 1.00 at the 1 per cent level of confidence. We find that six of the subtests have significantly inferior 15/50 ratios in Wechsler's population; all five of the performance subtests and the Digit Span. In the San Quentin sample, all the performance subtest 15/50 ratios are very significantly inferior, while two of the verbal ratios are significantly superior. The Spearman rank-order cor-

relation of the 15/50 ratios of these two populations is 0.97 which indicates that the relative order of dropping is about the same.

The differences in the various subtest 15/50 ratios for these two populations run from .02 to .34. The average ratio difference for Verbal is .18, while for Performance it is but .05. Figures 3 and 4 summarize these findings,

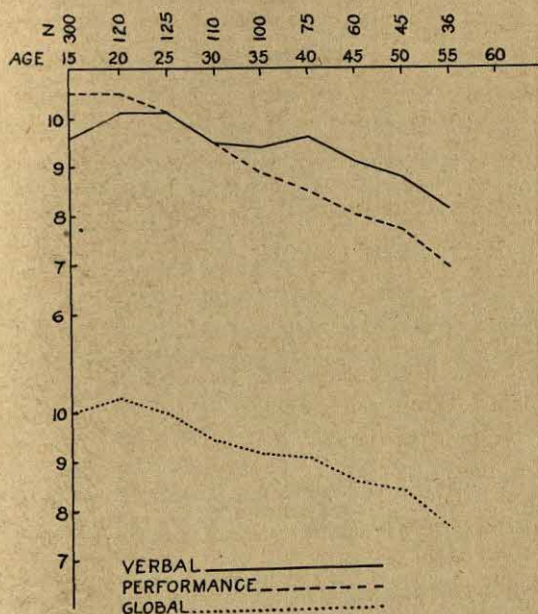


FIGURE 3

WECHSLER-BELLEVUE TEST, WECHSLER DATA: VERBAL AND PERFORMANCE SCORES
IN RELATION TO AGE

indicating the general trend lines for both populations. It can be seen that two entirely different curves are presented with respect to Verbal, while with respect to Performance, essential similarity is obtained. For the test as a whole, on the San Quentin population it appears that abilities hold rather well from 15 to about 55.

K. DISCUSSION

The important thing to bear in mind is not the similarities or the differences, but the fact that there are *both* agreements and disagreements. Were our findings similar to Wechsler, we could agree that his curves are probably general for the population as a whole, and then possibly that intelligence declines with greater maturity. Had the two sets of curves been strikingly

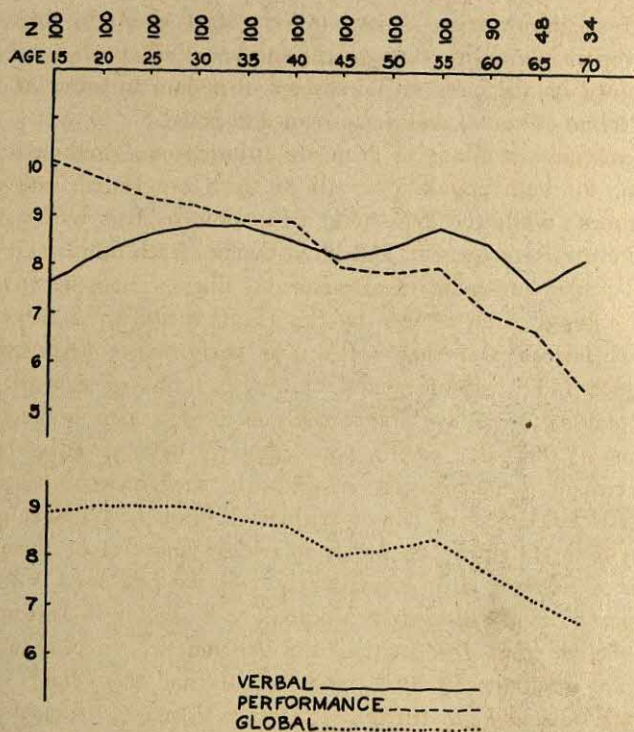


FIGURE 4
WECHSLER-BELLEVUE TEST, SAN QUENTIN DATA: VERBAL AND PERFORMANCE SCORES
IN RELATION TO AGE

different, all we would have known was that Wechsler's population and ours were different, but which population represented more closely the general rule would have been debatable. The fact is that in some respects the test results of the two populations are in agreement, while certain other results are not in agreement. This combination of agreements and disagreements raises questions which need to be resolved. The only identity in our two groups is age and the only agreements on the two sets of data are the Performance ratios. It seems then that the drop in Performance ability is a function of age, and not necessarily of intelligence.

In terms of verbal material, on the whole the Wechsler population showed a decline of 8 per cent while the San Quentin population showed an increase of 9 per cent. We doubt very much that this means that people increase in intelligence as they grow older, but it seems rather that they may add to

their store of information, or at the very least maintain their abilities in terms of verbal knowledge through the fifties. This conclusion could almost be maintained on the basis of Wechsler's own data in terms of the relative rates of decline of verbal and performance material.

The items most resistant to drop are Information, Arithmetic, and Comprehension, for both populations: all of them are verbal, and only Arithmetic is timed; while the items most susceptible to drop with age are Block Design, Picture Arrangement, and Digit Symbol, each timed, and each calling for visual acuity and motor performance. There cannot be any doubt that the rapid rates of drop as seen by the 15/50 ratios are a function of non-intellectual factors, and that verbal and performance items, while almost equally good indicators of mental ability in terms of general correlations with acceptable criteria, are differentially affected by non-intellectual matters.

In terms of the San Quentin population we have actually three types of subtests, two of them rise very significantly with increasing age: Information and Arithmetic; three remain within the limits of statistical expectation: Comprehension, Digit Span, and Similarities; and five of them, all of the performance subtests, drop significantly. If we had used Vocabulary, we could expect it too to rise, since Vocabulary is similar to Information. We would conclude that Information and Arithmetic are plus-culturally contaminated as intelligence subtests, and not only will they "hold" but they will "rise" with advancing maturity. It is easy to make a logical case for the increase for these subtests. The performance tests are minus-physiologically-contaminated due to visual, auditory, and motor demands. What remains is Comprehension, Similarities, and Digit Span as relatively uncontaminated subtests for the purpose of between-age research.

How can we explain the differences between Wechsler's population and the San Quentin population; how can we explain the differences of the subtest scores in the San Quentin population? And, how can we relate all of our findings to the general question of intelligence and aging?

We must agree that neither population can be proved to represent a "normal" population with respect to the important variable of intelligence. We suspect, despite the opinions of Rabin and Guertin, that Wechsler's normative population was unrepresentative, being too high at the younger end and too low at the older end. It may be that the San Quentin population is biased in the opposite direction.

In any event, the very similar rates of decline for all of the performance subtests for both populations would appear to be a function of aging, and not of intelligence. The fact that each of these subtests is a valid indicator

of intelligence within any age group does not vitiate the impression that as a "pure" measure of intellect for the purpose of between-age comparisons these tests are deficient, in that they will create the impression that 50-year-old people are only 75 per cent as intelligent as 15-year-old people, which we know cannot be true on the basis of the relative performance of older people in verbal tests. The older group drops only 8 per cent on Verbal in Wechsler's data. This could very easily be a function of poor sampling. The San Quentin Verbal rose 9 per cent. If this is also a function of poor sampling, the consensus would be parity, that is to say that as far as the verbal items measure intelligence, there is no difference in the abilities of 15- and 50-year-old people.

When we examine our individual subtests we are struck with the fact that in both populations, the rank order of "drop" is $\rho = .97$. The two populations appear to be operating in the same relative manner with respect to decline. But the 15/50 ratio difference of 56 between Information and Digit Symbol for our population is so excessive as to create the definite impression that both of these subtests are contaminated by additional cultural experience in the one case and by visual-motor losses in the other.

We believe that only Comprehension, Similarities, and Digit Span, of the 10 Wechsler subtests, can be used for between-age research. The reasons that these three are relatively unaffected are not clear, and we would be guilty of post-facto reasoning were we to try to justify their relative purity. The Wechsler-Bellevue Test consists of 11 subtests, of which only 10 are under discussion. There is no reason why this test could not have included many more subtests, nor why it might not have had fewer. If the *W-B* test were to have consisted of Information, Comprehension, Arithmetic, and Similarities, our conclusions in reference to intelligence and aging, from Wechsler's own data, would have been that there is no significant drop. On the other hand, if the *W-B* test had consisted of the five performance tests, then we would have concluded that intelligence declined significantly with maturity.

If the *W-B* test had consisted of Information and Arithmetic, on the basis of our data, we would have concluded that people increase in intelligence as they grow older; but if the test consisted of the performance subtests, we would have concluded that intelligence declined.

The most sensible conclusion is to declare that subtests are affected by various non-intellectual factors, such as greater experience, education, visual loss, motor loss, and that as people grow older these peripheral or non-pertinent factors act as artifacts which tend to distort test results. Performance tests are clearly unsuited for the purpose of between-age research in intelli-

gence. Verbal tests which include cultural material of the type that is constantly added to as one grows older are also unsuitable for a diametrically opposed reason.

When we use various combinations of subtests we can not reach useful conclusions unless we have an equal and opposing number of subtests compensating for plus and minus errors.

In view of these findings and arguments, it would seem most scientific to conclude that, in the matter of increasing age not causing a decline in intelligence, the null hypothesis has not been disproved.

L. SUMMARY

Although there appears to be no disagreement among investigators that intelligence test raw scores decline with advancing age, this fact is variously interpreted by various authorities, some believing that intelligence remains constant to extreme old age and some believing that it declines following early maturity.

Examining four large-scale investigations of intelligence and aging critically for sampling and technique of measurement, we come to the conclusion that proper sampling is not proven for any of the investigations and that the use of group tests in three cases was inappropriate.

We argue for the representatives of a forced sampling such as ours in prison and we present data on 1,072 adults, with 372 above the age of 49, using the Wechsler-Bellevue Test.

We find essential similarity in terms of rates of decline between our sample and Wechsler's population for performance material, but significant differences between Wechsler's decline on Verbal and Performance, as well for our sample. We further find significant differences between our and Wechsler's Verbal data. Our conclusions follow:

1. General intelligence does not decline from early to late maturity.
2. Test ability will vary in a downward direction if the subtests contain visual and motor factors, and will vary in an upward direction if the subtests contain material which depends on continued learning.

REFERENCES

1. BENSBERG, G. L., & SLOAN, W. A study of Wechsler's concept of mental deterioration in older mental defectives. *J. Clin. Psychol.*, 1950, **6**, 359-362.
2. DRESSLER, D. Probation and Parole. New York: Columbia Univ. Press, 1951.
3. JONES, H. E., & CONRAD, H. S. The growth and decline of intelligence. *Genet. Psychol. Monog.*, 1933, **13**, 223-298.
4. KUHLEN, R. G. Psychological trends and problems in later maturity. In L. A. Pennington and I. A. Berg (Ed.). *An Introduction to Clinical Psychology*. New York: Ronald Press, 1948.

5. MILES, C. C., & MILES, W. R. The correlation of intelligence scores and chronological age from early to late maturity. *Amer. J. Psychol.* 1932, **44**, 44-79.
6. MURCHISON, C. A. Criminal Intelligence. Worcester, Mass.: Clark Univ. Press, 1926.
7. RABIN, A. I., & GUERTIN, W. H. Research with the Wechsler-Bellevue Test. *Psychol. Bull.*, 1951, **48**, 211-248.
8. ROUKE, F. L. Recent contribution of psychology to the study of criminogenesis. *J. Crim. Law & Criminol.*, 1950, **41**, 446-455.
9. SCHMIDT, D. G. Levels of intelligence of prison inmates. *Amer. J. Ment. Def.*, 1946, **51**, 63-66.
10. SPEARMAN, C. The Abilities of Man. New York: Macmillan, 1927.
11. SUTHERLAND, E. H. Mental deficiency and crime. In K. Young (Ed.). *Social Attitudes*. New York: Holt, 1931.
12. TULCHIN, S. Intelligence and Crime. Chicago: Univ. Chicago Press, 1939.
13. WECHSLER, D. The Measurement of Adult Intelligence. Baltimore: Williams & Wilkins, 1944.
14. WEISENBERG, T., ROE, A., & MCBRIDE, K. E. Adult Intelligence. New York: Commonwealth Fund, 1936.
15. YERKES, R. M. Psychological Examining in the U. S. Army. *Mem. Nat. Acad. Sci.*, **15**, 1921.

5306 S. Greenwood Street
Chicago 15, Illinois

THE DOMINANT METHOD OF IMAGERY IN BLIND AS COMPARED TO SIGHTED ADOLESCENTS*

Department of Ophthalmology, Indiana University School of Medicine

T. F. SCHLAEGEL, JR.¹

A. THE PROBLEM

There are various methods of imagery. For example, if several people are asked for their mental image of a "crying baby" there will be differences in the images they experience. Some will see the baby; others will hear it; some individuals will smell the baby or feel the touch of its skin, or will notice whether it is hot or cold; whereas others will notice their muscles work to pick it up. Usually more than one method is employed. The modalities considered in this study are: visual, auditory, kinesthetic, tactile-temperature, and olfactory-gustatory.

In 1838, Heermann (5) and in 1888, Jastrow (6) ascertained by statistical inquiry among the totally blind that if their blindness occurred before a period embraced between the fifth and seventh years, visual dreams were gradually outgrown. If sight was lost after the seventh year, visual dreams were found to survive throughout life.

In 1914, Town compared six high school subjects blind since infancy with six normal adults. She spelled words backwards to them and asked them to determine the rôle of visualization in naming the words (9). She concluded that the test had little or no value in determining the rôle of visual imagery. In 1932, Bowers (1) studied the visual imagery of sighted adolescents. The subjects were asked to look at a clear photograph and then to compare the visual image they were later able to elicit with a series of more and more blurred copies of this photograph. He concluded that his test results were independent of (a) age, within the range of 12-21 years, (b) sex, within the range of 12-21 years, (c) *IQ*, and (d) ability in certain school subjects.

The purpose of the present imagery study of blind school subjects was to determine the rôle of two factors: 1. visual acuity, 2. age of onset of incapac-

*Received in the Editorial Office on October 9, 1951.

¹The author wishes to express his appreciation to Roy Adams for his advice and assistance in the treatment of the statistical data of this study.

tating visual loss. Both of these factors proved to play an important rôle in the imagery employed.

B. MATERIALS AND METHODS

To determine the dominant method of imagery, a modification of Griffitt's test (4) of "concrete imagery" was used. As adapted to this purpose, the test consisted of 125 words or phrases to which the subject responded with his *initial* mental image. Samples of the test words used are: dog, street-car, George Washington. Samples of the test phrases used are: a girl talking on a telephone; a train pulling out of station; the driver whipping the horses. This test was given to blind school students and to a sighted group.

1. *Blind Group*

At the Indiana School for the Blind, 67 pupils ranging in age from 12 through 24 years with an average of 16.2 years were assembled and supplied with braille writing equipment. The author explained the test and made sure the subjects understood it by the use of trial samples. They were asked to write down one of the following words as an answer: see, hear, muscle, touch, temperature, smell, or taste. After questions were answered and it was felt that the test was understood, the examiner called out the words and phrases one at a time and the subjects wrote their first mental image in braille. These braille answers were translated by the faculty of the Blind School.

At a separate period, each of the Blind School subjects was interviewed concerning the date of onset of his significant visual loss.

The 67 Blind School subjects represent a heterogeneous group, not only in regard to age of onset but also in regard to the degree of vision. The visual acuity of the Blind School subjects was determined. On the basis of present vision, they were divided into three groups: (a) 0 vision: absolute blindness up to light perception only; (b) \pm vision: the ability to detect any movement or objects to counting fingers at 5 feet; (c) + vision: vision better than 5/200.²

2. *Control (Sighted) Group*

At Shortridge High School, Indianapolis, 78 high school students, with ages ranging from 14 through 18 years and an average age of 15.4, were assembled. They were supplied with one sheet of paper numbered from 1 to 125.

²There was a close correlation between the distance at which fingers were counted and the distance at which the 20/200 E was seen, i.e., CF at 5 feet is roughly equal to 5/200.

The test was given to this control group by the same examiner, and an attempt was made to duplicate the conditions of the test as given to the blind subjects. The test was explained, examples tried, and the test words and phrases called out while the subjects wrote down one of the following seven answers: see, hear, muscle, touch, temp., smell, or taste.

It was considered unnecessary to check the vision of the Shortridge subjects, since one of the main criteria for admission to the School for the Blind is inability to attend other public schools because of poor vision.

C. DATA AND RESULTS

The raw data for the Blind School group is presented in Table 1 and that for the sighted group in Table 2. Because of their small number and similarity the tactile and temperature responses have been grouped together as have the olfactory and gustatory. The one significant difference between the averages of these two groups lies in the number of unknown responses. This unknown group included: those answers which we were unable to decipher; answers not one of the seven acceptable ones and incapable of being translated into one of them; and failure to answer. The greater number of these unknown answers in the Blind School group is considered to be due mainly to difficulty in the use of braille. Many of the subjects at the Blind School had not been there long enough to achieve complete mastery of its use.

1. *The Dominant Method of Imagery*

Definitions of dominance that could be used are: (a) A majority of responses of one type (63 or more). (b) One-third, or more, responses of one type (42+). (c) The type of imagery used most frequently.

Calculations comparing these three methods of defining dominance showed that the last definition gave the best spread of cases and that the other definitions did not add to the understanding of the problem.

When the averages of the blind and sighted groups are considered as a whole the relative dominance is the same for both in the following decreasing order; visual, auditory, kinesthetic, tactile-temperature, and olfactory-gustatory (Table 1).

2. *The Rôle of Present Visual Acuity in the Imagery Employed*

Table 3 illustrates that there were significant differences between the three groups divided on the basis of present vision. (a) The Blind School subjects with the best vision (+) rated significantly higher in the percen-

TABLE 1

The number of visual, auditory, kinesthetic, tactile-temperature, olfactory-gustatory, and unknown imagery responses out of a total of 125 for each of 67 blind school subjects with their age, visual acuity, visual class, and age of onset of incapacitating visual loss. For quick comparison the percentages of both the blind and sighted control are included at the bottom of the table.

Vis	Method of imagery					Ages	Visual acuity	Visual class	Age of onset
	Aud	Kin	T-T	O-G	Unk				
44	40	13	20	5	3	13	CF 3 ft.	±	Birth
37	29	33	17	8	1	14	No L.P.	0	6 years
83	27	0	3	1	11	18	5/70-1	+	Birth
115	9	1	0	0	0	17	5/10	+	Birth
21*	56	10	17	15	6	18	No L.P.	0	Birth
40*	43	22	8	10	2	20	L.P.	0	Birth
34	66	11	7	7	0	16	No. L.P.	0	5 years
53	30	19	15	8	0	20	No L.P.	0	8 years
40	42	20	14	9	0	17	L.P.	0	10 years
115	3	2	1	1	3	14	5/100	+	Birth
30	24	12	12	3	44	15	CF 2 ft.	±	8 years
101	23	0	0	0	1	14	5/70	+	5 years
68	41	5	4	6	1	16	5/30	+	14 years
1	49	29	25	12	9	20	No L.P.	0	Birth
44	40	26	7	6	2	15	5/100	+	Birth
22	52	21	20	10	0	17	Poor L.P.	0	13 years
55	50	5	9	6	0	15	5/70	+	Birth
7*	72	10	26	9	1	19	No L.P.	0	Birth
80	37	2	1	2	3	16	5/70-1	+	Birth
87	24	3	4	6	1	19	CF 5 ft.	±	6 years
89	27	0	0	1	8	24	5/70	+	11 years
34	43	27	8	9	4	17	HM 6 ft.	±	Birth
53	35	15	9	9	4	19	No L.P.	0	9 years
82	29	9	2	2	1	21	HM 5 ft.	±	10 years
19	44	22	24	6	10	19	HM 1 ft.	±	Birth
57	45	6	9	7	1	17	12/200	+	Birth
2	75	17	21	9	1	20	L.P.	0	Birth
78	11	20	11	5	0	16	No L.P.	0	10 years
105	12	4	1	2	1	14	CF 4 ft.	±	11 years
32	40	32	13	8	0	18	L.P.	0	7 years
37	40	32	8	7	1	16	5/70	+	Birth
108	6	1	5	3	2	20	HM 3 ft.	±	11 years
52	51	6	8	8	0	17	HM 1 ft.	±	Birth
0	70	6	35	10	4	17	No L.P.	0	8 years
56	46	12	3	6	2	17	5/100	+	Birth
85	27	2	1	9	1	18	5/70	+	10 years
49	47	4	15	8	2	15	HM 4 ft.	±	Birth
47	51	8	5	5	9	13	L.P.	0	7 years
31*	68	7	10	9	0	12	L.P.	0	5 years
10	59	28	19	9	0	24	HM 6 ft.	±	Birth
90	17	5	2	2	9	19	L.P.	0	16 years
31	39	10	12	9	24	15	CF 5 ft.	±	1½ years
82	40	0	0	3	0	14	5/50	+	Birth
103	15	4	2	1	0	15	5/50	+	Birth
63	40	6	4	5	7	15	5/70	+	Birth
49	47	11	7	11	0	16	5/100	+	6 years
27	38	32	15	13	0	14	5/50	+	1 year

TABLE 1 (continued)

	Vis	Aud	Method of imagery				Unk	Ages	Visual acuity	Visual class	Age of onset
			Kin	T-T	O-G						
	42*	55	11	13	4	0	12		No L.P.	0	2 years
	53	44	5	13	9	1	12		5/15	+	7 years
	73	35	4	7	5	1	13		5/70	+	Birth
	64	41	10	5	5	0	13		CF 5 ft.	±	Birth
	54	43	11	9	8	0	20		CF 1 ft.	±	7 years
	53	56	2	4	4	6	12		HM 2 ft.	±	9½ years
	61	37	3	3	4	17	13		CF 4 ft.	±	4 years
	0	42	48	21	7	7	12		No L.P.	0	Birth
	24	25	24	29	20	3	14		HM 2 ft.	±	Birth
	49	43	12	13	3	5	12		CF 5 ft.	±	1½ years
	112	12	0	1	0	0	16		HM 1 ft.	±	Birth
	112	4	6	1	1	1	16		No. L.P.	0	9 years
	28*	36	31	18	11	1	21		L.P.	0	Birth
	30*	41	28	18	3	5	13		No L.P.	0	Birth
	34*	46	11	20	9	5	19		No L.P.	0	Birth
	0	62	16	33	10	4	19		HM 8 ft.	±	Birth
	8*	58	40	9	8	2	15		Poor L.P.	0	Birth
	58	40	10	7	10	0	17		Cf 5 ft.	±	6 years
	59	30	5	8	4	19	13		9/200	±	Birth
	117	6	0	0	2	0	12		5/50	+	8 years
Total	3,579	2,575	847	691	427	256	1,089			—	—
Mean			53.4	38.4	12.6	10.3	6.4		3.8 16.2	—	3.8
Percentage				42.7	30.7	10.2	8.3		5.1 3.0	—	—
Percentage for 78											
Normal Subjects											
from Table 2				41.7	30.8	10.2	9.8		7.1 0.4	—	—

*Although these subjects answered, "see," it was discovered by interview that they just "thought of" the object.

tage of responses of the visual type. Fifty-eight per cent of their imagery responses were of the visual type as compared with 41 per cent for the sighted and 42 per cent for all blind school subjects. (b) The average number of responses of those subjects with intermediate vision (\pm) did not differ significantly from those of the control group or from those of the entire Blind School group. (c) Those subjects with the poorest vision (0) had the lowest average number of visual responses and the highest average number of auditory responses. This group alone had auditory dominance.

Table 4 depicts this information in a slightly different form. Instead of comparing the average percentage scores, the dominant modality for each individual is considered. Each score in the table represents the number of subjects using that modality as their dominant method. Again, the importance of audition in those with 0 vision is seen in the fact that 70 per cent (16 of 23) had auditory imagery as their dominant method.

TABLE 2

The number of visual, auditory, kinesthetic, tactile-temperature, olfactory-gustatory, and unknown imagery responses out of a total of 125 for each of 78 Shortridge High School sighted adolescents, plus the age of each subject.

Vis	Method of Imagery				Unk	Age
	Aud	Kin	T-T	O-G		
33	52	23	6	11	0	17
35	38	11	28	13	0	17
37	57	9	11	9	0	15
38	36	25	16	10	0	15
52	26	18	18	11	0	15
51	38	14	13	8	1	15
60	31	12	15	7	0	15
35	28	25	25	12	0	15
62	29	8	22	4	0	16
50	44	8	13	10	0	15
50	57	2	7	8	1	15
52	33	5	24	10	1	16
83	25	9	1	7	0	15
33	56	13	13	10	0	17
48	40	12	16	8	1	15
43	42	12	16	12	0	15
57	38	13	8	8	1	15
64	40	4	6	10	1	15
70	35	6	7	7	0	16
31	31	13	37	12	1	15
43	27	35	10	10	0	16
44	62	15	6	8	0	16
32	49	13	17	12	2	16
25	49	10	21	20	0	16
89	33	0	0	3	0	15
68	28	14	7	8	0	15
40	43	26	4	12	0	16
66	37	4	10	5	3	15
48	48	7	13	8	1	17
51	51	6	9	7	1	15
49	36	16	16	8	0	15
83	22	8	9	3	0	14
53	37	13	14	8	0	15
42	50	10	11	12	0	15
81	23	10	7	4	0	16
49	43	9	8	13	3	15
46	51	17	4	6	1	15
54	27	15	17	8	4	15
30	33	14	26	22	0	15
39	50	10	14	10	2	15
52	45	9	11	8	0	15
60	34	6	12	9	4	15
61	39	13	7	5	0	16
25	49	23	15	13	0	15
49	47	17	5	7	0	15
50	53	6	5	11	0	16
42	46	10	18	9	0	16
31	31	25	30	8	0	15
40	50	22	4	9	0	16
58	33	15	8	9	2	15

TABLE 2 (continued)

	Vis	Method of Imagery				Unk	Age
		Aud	Kin	T-T	O-G		
	35	41	17	18	14	0	15
	65	45	5	5	5	0	16
	45	45	20	5	10	0	15
	33	52	15	13	12	0	15
	39	38	22	16	9	1	15
	31	45	18	20	11	0	17
	57	40	5	13	10	0	15
	47	40	12	16	10	0	15
	54	39	16	7	9	0	15
	62	48	4	2	8	1	15
	61	45	3	7	9	0	14
	61	27	19	10	8	0	15
	50	26	19	21	9	0	15
	46	39	11	16	12	0	15
	93	18	7	2	4	1	15
	95	22	4	3	1	0	15
	68	14	29	6	8	0	15
	46	35	13	19	12	0	15
	59	32	11	11	12	0	16
	36	31	21	29	7	1	16
	93	18	8	5	1	0	15
	63	40	7	8	7	0	16
	49	49	14	5	8	0	16
	32	37	22	19	13	2	15
	85	26	4	4	6	0	18
	43	48	16	10	7	0	15
	74	28	9	4	8	2	16
	55	43	7	12	8	0	15
Total	4,061	3,003	998	948	700	40	1,200
Mean	52.1	38.5	12.8	12.2	8.9	0.5	15.4
Percentage	41.7	30.8	10.2	9.8	7.1	0.4	—
Percentage for blind school subjects							
from Table 1	42.7	30.7	10.2	8.3	5.1	3.0	—

TABLE 3

The average percentage of visual, auditory, kinesthetic, tactile-temperature, olfactory-gustatory, and unknown imagery responses for 67 Blind School subjects divided according to their degree of vision and compared with the sighted control.

Type of imagery	Sighted control	0 Vision (No L.P. to L.P.)	± Vision (H.M. to CF at 5 gt.)	+ Vision better than 5/200
Vis.	41.7	27.9	42.4	58.6
Aud.	30.8	36.4	29.8	25.8
Kin.	10.2	15.3	8.9	5.8
T-T	9.8	11.9	9.0	3.7
O-G	7.1	6.2	5.2	3.8
Unk.	0.4	2.3	4.7	2.3
Total %	100.0	100.0	100.0	100.0

TABLE 4

The dominant method of imagery of each of the 67 Blind School subjects scored under the various methods of imagery and classed according to the degree of vision.

Type of imagery	0 Vision (No L.P. to L.P.)	\pm Vision (H.M. to C.F. at 5 ft.)	+ Vision (from 5/200 up)
Vis.	6	13	20
Aud.	16	7	2
Kin.	1 •	0	0
T-T	0	1	0
Unk.	0	1	0
Total	23	22	22

$P < .05$ —distribution significant.

3. *The Rôle of the Age of Onset of Incapacitating Visual Loss*

Figure 1 illustrates that the age of onset played the greatest rôle in the amount of visual imagery employed in those Blind School subjects with 0 vision, but as the degree of vision increased, the age of onset was less im-

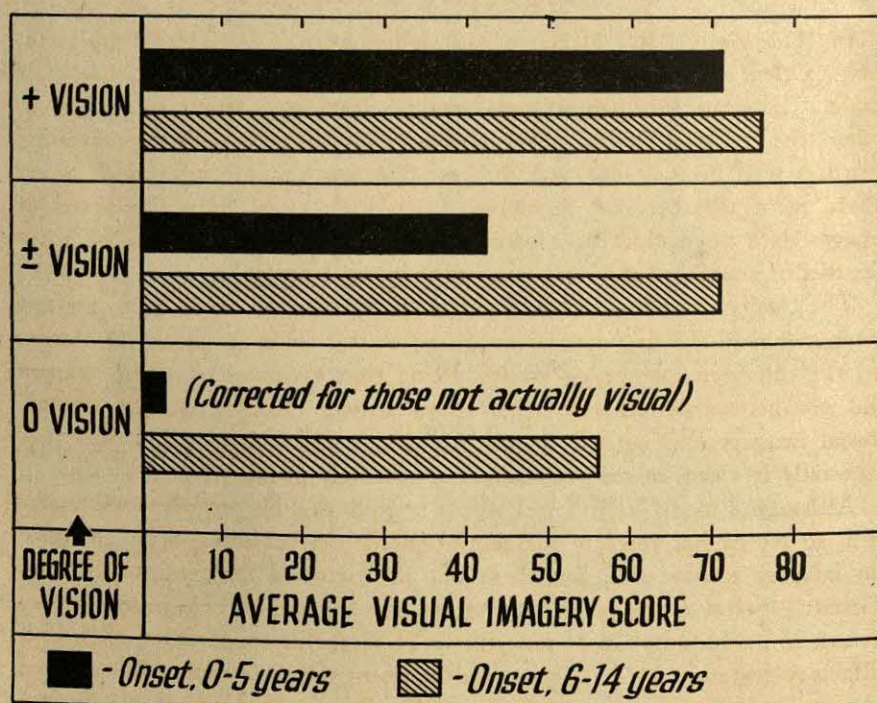


FIGURE 1

AVERAGE VISUAL IMAGERY SCORES IN BLIND SCHOOL SUBJECTS AS AFFECTED BY VISION AND AGE OF ONSET OF "BLINDNESS"

TABLE 5

The dominant method of imagery of each of 23 Blind School subjects with 0-Vision grouped according to whether the onset of their incapacitating visual loss was before or after 6 years of age.

Age of onset of blindness	Number of subjects	Dominant method of imagery		
		Visual	Auditory	Kinesthetic
0-5 years	13	0	12	1
6-14 years	10	6	4	0

$P < .01$ —distribution significant.

portant. In Table 5 it is seen that none of the 0-vision students who lost their sight before the age of 6 had visual imagery dominant (only three had *any* visual imagery): whereas among those who lost their vision later, 60 per cent (6 of 10) had vision as their dominant method of imagery.

D. COMMENT

1. *"Visual Imagery" in the Sighted*

In 1883 Galton (3) asked various sighted persons to describe the vividness of their mental picture when calling to mind the morning's breakfast table. To some the mental scene was as clear and natural as reality; to others the resulting mental image was tolerably distinct, with the prevailing features well brought out, but the rest dim and vague; and a third group could piece together only a vague, fragmentary, and unreliable series of images with no distinct or constant picture. Thus Galton described three degrees of sharpness of visual imagery in normal people.

The clarity of mental imagery played no rôle in this study, since the test requested only the first mental image no matter what its degree of clarity. In the interviews, however, it was found that vividness of visual imagery did not necessarily decrease with visual acuity. In the partially sighted, visual imagery may actually be clearer than present vision. This was true especially if vision in the past had been relatively good.

Although I used Griffitt's test words and phrases, his procedure was modified, so we cannot refer to his results on normal subjects. Seitz (8) gave an imagery test as was done here but he borrowed different words from Griffitt. Seitz' mean percentage scores for a group of 114 nurses were: visual 25.2; auditory 23.2; kinesthetic 21.5; tactile-temperature 15.9; and olfactory-gustatory 14.2. These results present the same order of dominance obtained in the present study. In 1947, Bower (1) studied the images aroused by reference to onions frying in a pan and obtained the following order of diminishing frequency: visual, auditory, tactile, kinesthetic, thermal,

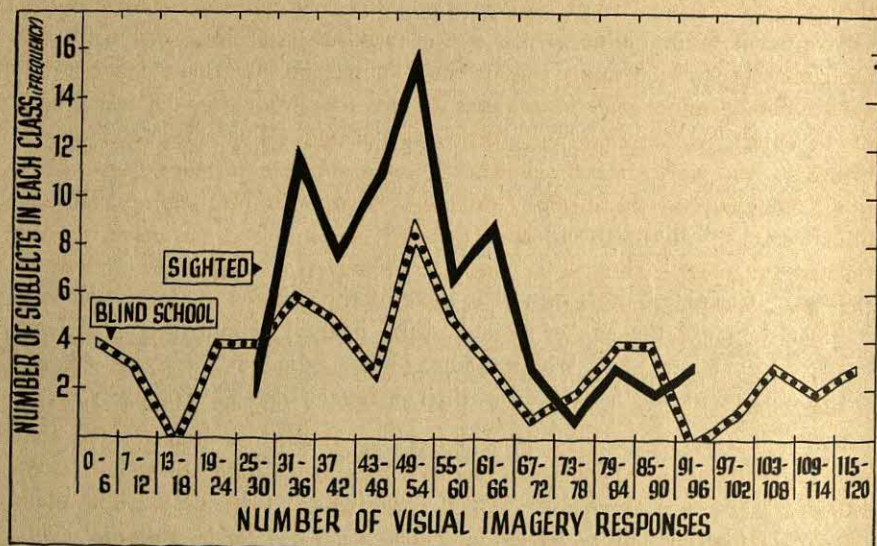


FIGURE 2
RANGE AND FREQUENCY OF VISUAL IMAGERY RESPONSES

olfactory. It is apparent that the content of the stimulus will have an effect on the results obtained; nevertheless, there is a striking similarity in the order of dominance obtained by various investigators.

One source of possible error in the present experiment lies in the greater range of ages in the Blind School group (12-24 years) as compared with the control (14-18 years), but in Bower's study (1) age, in the 12-21 period, played no rôle in his visual imagery results.

Setting a date for the onset of significant visual loss was difficult in many cases. Most of the subjects had more than light perception and in most cases the vision had not been stationary but had decreased over the years. The final judgment was made after consulting the medical and parental records and after talking with the subjects.

2. "Visual Imagery" in Those Blind Since Birth

An unexpected finding was the number of visual answers in those blind since birth. Because of this result, the students were interviewed. In talking with them, it was decided that what they meant by "see" was not a visual image but that they thought of the characteristics of the object named. For example, in response to the test words "George Washington" they thought to

themselves, "He is tall with a large frame, has gray hair, and a large nose." Of course, it is impossible to prove that these subjects blind since birth did not construct a visual image out of their knowledge, but this seems improbable. The examiner and the subjects decided that what they did was "think of the object" rather than perceive a visual image. This impression is confirmed by the work of Librachowa (9) who obtained introspective descriptions and analyzed the mental "products" from over 600 subjects plus 20 blind persons. In the blind she frequently encountered the purely verbal type of statement. Because of the lack of the visual factor the content of the idea appeared exclusively.

These responses which were apparently not visual were left under the visual column, however, since calculating them separately produced no change in the significance of the results. Two of the 10 subjects blind since birth actually had some meager visual imagery. It will be remembered that under the 0 class of vision were included those with light perception. One person has one visual response which consisted of colors only. The other person had two visual responses in which she noted lights flashing.

3. *The Rôle of Vision in the Choice of the Dominant Method of Imagery*

Although the averages of the Blind School group compare closely with those of the sighted group, there is a greater variability among the children at the Blind School. This is seen in the wider range of responses (Tables 1 and 2, Figure 1). Since the Blind School subjects are not homogeneous with respect to vision, it was postulated that this variability of vision might explain the wider range of responses.

Both from common sense and the work of others we would have expected that visual imagery would not be dominant among the truly blind. The one unexpected result of this study was the high percentage of visual imagery in the semi-blind (+ and \pm vision). Those with \pm vision had visual imagery scores equal to the sighted and those with + vision had scores significantly higher. Why did these semi-blind subjects have more than the expected amount of visual imagery? Because of the importance of vision, the child may respond to the pressure of this importance by compensatory visual imagery. This question requires further investigation. It would be interesting to extend this study of imagery to cover the complete range of visual acuity from absolute blindness to 20/10 vision. Because of individual variation a large number of subjects would be needed.

4. *The Rôle of the Age of Onset of Incapacitating Visual Loss on Visual Imagery*

Heerman (5) and Jastrow (6) found in studying visual dreams in the blind that the age of onset of blindness was important. The critical years were from 5 to 7. Those who lost their vision before these ages had no visual dreams. When blindness had occurred between the age of 5 and 7, some did and some did not experience visual dreams. In this study we have taken the age of 6 as the dividing line, but our data more closely support the idea of a transition zone. Of the 13 subjects with 0 vision blinded before the age of 6, only three had any visual imagery. Two of these had had light perception since birth. One had one visual response of colors only and the other subject had two visual responses of lights flashing. The third case with visual imagery in this group had had his onset at the age of 5 and scored 34 visual responses. There was one subject with loss of vision at the age of 6 who had 37 visual imagery responses. Another one, a girl completely blind since the age of 8, had no visual responses. Thus our results indicate a transition zone from 5 through 8 years of age, since when blindness had occurred during this period, some did and some did not possess visual imagery. An interesting question is why those who lose their sight before this transition period lose their ability for visual imagery and dreams. Is it disuse from lack of reinforcement or is it due to repression? An answer to this question might be obtained by regression under hypnosis to check for the presence or absence of visual imagery at an early age.

E. SUMMARY

An imagery test was given to 67 Blind School, and to 78 sighted, adolescents. When the averages of the blind and sighted were considered as a whole they were essentially equal and the relative dominance was the same for both in the following decreasing order: visual, auditory, kinesthetic, tactile-temperature, and olfactory-gustatory. However, when the Blind School subjects were divided on the basis of present vision into three groups, significant differences in their imagery were found. When each of these three groups was divided into two subgroups on the basis of whether the age of onset of incapacitating visual loss was before or after 6 years of age, further differences were found.

It was thus found that the imagery of Blind School subjects was significantly affected by two factors.

1. *Present visual acuity.* Those subjects with the poorest vision had the

least number of visual, and the greatest number of auditory, responses. As the visual acuity (at Blind School levels) increased, there was an average increase in visual imagery responses to an extent even greater than that of normal controls.

2. *Age of onset of incapacitating loss of vision.* If the onset was before the age of six, visual imagery tended to disappear, being most pronounced in those subjects with the poorest vision.

REFERENCES

1. BOWER, H. Visual imagery of high school pupils. *J. Genet. Psychol.*, 1932, **40**, 232-234.
2. BROWER, D. The experimental study of imagery: II. The relative predominance of various imagery modalities. *J. Gen. Psychol.*, 1947, **37**, 199-200.
3. GALTON, F. *Inquiries into Human Faculty*. London: Macmillan, 1883.
4. GRIFFITTS, C. H. *Fundamentals of Vocational Psychology*. New York: Macmillan, 1924.
5. HEERMANN, G. Beobachtungen und Betrachtungen über die Traume der Blinden, ein Beitrag zur Physiologie der sinne. *Monat. Med. Augenh.*, 1838, **3**, 116-180.
6. JASTROW, J. The dreams of the blind. *New Princeton Rev.*, 1888, **5**, 18-34.
7. LIBRACHOWA, M. Structure of individual (specific) reproductive images and generic images. *Kwart. Psychol.*, 1934, **5**, 124.
8. SEITZ, P. F. D. Relation of mental imagery to hallucinations. *Arch. Neurol. & Psychiat.*, 1947, **57**, 469-480.
9. TOWN, C. H. The contribution of visual imagery to verbal thought: A comparative study of seeing and blind subjects. *Illinois Med. J.*, 1914, **26**, 354-358.

Department of Ophthalmology
Indiana University School of Medicine
1040 West Michigan Street
Indianapolis 7, Indiana

STUDIES IN DELINQUENCY: III. AN INDIVIDUAL TEST*

Teachers College of Connecticut, New Britain, Connecticut

F. C. ZAKOLSKI

A. PROBLEM

In the reported studies on delinquency in boys (6), particularly that dealing with prediction (7), certain desiderata remained: (a) Statistical aspects presented problems inadequately solved. An increase in the number of cases would insure to a larger degree the adequacy of the sample; it would reduce the magnitude of the errors calculated. Determination of the errors of the percentiles was necessary to indicate the possible variability and to demonstrate the probable position of the true percentiles in order to avoid the impression of an inflexibly determined score. Errors of percentiles also vary in magnitude in accordance with the magnitude of the percentile; this should be indicated. The extent to which percentages within intervals distinguished the two groups was unknown. Further, the method of ascertaining the validity of the test, by the bi-serial coefficient of correlation, failed to indicate the great amount of overlapping between delinquent and non-delinquent scores. Also certain reliabilities, for example the test-retest reliability, were undetermined. (b) The items presented problems. Some items were ineffective; some were unnecessarily difficult to read, particularly for delinquent groups who, as a group, are generally not academically minded; the final reading difficulty of the items was also unknown. (c) Also in need of change was the form of the test and the method of analyzing and reporting the results, and reconstructing the personality pattern.

B. PROCEDURE

Items were prepared. Some items were unnecessarily difficult because of the words involved. Substitutions were effected: "frequently" became "often," in itself a somewhat difficult word though easier than the former so that the problem was only mitigated; "rapidly" became "fast"; "pleasant" became "good," although the connotation was influenced. In some cases, it was not possible to find a synonym; "hypnotizing" was one of these; "mesmerizing" increased the difficulty. The denotation of "hypnotizing" was known with great frequency but the word is difficult to read. Such words

*Received in the Editorial Office on October 16, 1951.

did not present an impasse since any of the meanings may be explained by the clinician, but they do make the administration of the test less convenient than it would be without them. When the items were simplified as much as possible, their reading difficulty was determined by the Dale formula (3). It was found to be 4.2; those with a reading ability equivalent to that of pupils in the second month of the fourth grade could read the items fairly successfully.

Responses of 50 delinquent and 50 non-delinquent boys were used for item selection. P was read directly from Zubin's unpublished monograph (Chart III) with the help of those published (8), and the level $P = 0.01$ was chosen as indicating true differences between the two groups. The equivalent values of P in units of normal distribution were used as a basis for unitary weights as shown in Table 1.

TABLE 1 (4)

P	X	Difference	Weight
0.01	2.3263		1
0.005	2.5768	0.2495	2
0.001	3.0902	0.5144	3
0.0005	3.2905	0.2003	4
10^{-4}	3.7190	0.4285	5
10^{-5}	4.2644	0.5458	6
10^{-6}	4.7534	0.4886	7
10^{-7}	5.1993	0.4459	8
10^{-8}	5.6120	0.4127	9
10^{-9}	5.9978	0.3858	10

When the items were weighted, each item was matched with another item of equal weight. The method of logically equivalent rather than random split halves seemed defensible for the calculation of odd-even reliabilities particularly because of the heavy weighting random splitting would have yielded, probably, smaller rather than truer reliabilities (2). This last result could be expected from a system which balanced a weight of "one" against a weight of "ten." By the method of logically equivalent halves, the half composed of odd items became theoretically equal to the half formed by even items; the two halves appeared, and mathematically were, equivalent. This procedure, however, has the disadvantage of sacrificing some items because they cannot be matched statistically. Determination of item validity also has an important relationship to test validity.

The form of the test was changed to an individual one. The items were placed on cards to be sorted into "yes" and "no" boxes. This arrangement is cumbersome but, it was hypothesized, would give a greater sense of

anonymity and so better rapport and franker responses. To this form, the individual does not have to sign his name, a particular advantage in dealing psychologically with delinquent, and probably non-delinquent, boys. The more complex procedure may also serve as a bulwark against irresponsible usage.

A residuum of 730 white boys served as subjects. The actual number was somewhat greater by approximately 70, but the data yielded were not found usable in every instance for various reasons, such as the presence of delinquent boys in supposedly non-delinquent groups, or because of the failure to gain rapport as indicated by the truth score. Sometimes resistance stemmed not from concern over delinquent behavior but from fears having a different origin. One individual, a non-delinquent, would not reveal himself at all; he was subject to severe epileptic seizures and was afraid that the method would reveal this. The groups were matched approximately for age. The range for both groups was 9-17 years. These years are often crucial in determining delinquent trend. The mean of the 365 non-delinquents was 12.3 ± 1.8 ; the mean for the delinquents was 12.5 ± 1.7 . The non-delinquent subjects were listed in public schools: Safford Junior High School, Tucson, Arizona, and the Central Junior High School, New Britain, Connecticut. Delinquent subjects were studied at the following locations: the Fort Grant Industrial School, Fort Grant, Arizona; the Fred C. Nelles School for Boys, Whittier, California; the Connecticut School for Boys, Meriden, Connecticut.

C. RESULTS

1. Means and Errors

Some pertinent statistical determinations are shown in Table 2. The former study (7) stated the mean (M) for the delinquents (D), when the cases (N) are 265, as 327.0, with an uncalculated error; the standard error (σ), or sigma, for the delinquents was 90.5 M for non-delinquents (ND), when $N = 265$, is 126.0, with a sigma of 58.1. The present study yields an M , for D , of 313.5 ± 4.81 , when $N = 365$; and sigma as 91.8 ± 2.40 . For ND , where $M = 128.4 \pm 3.13$; sigma is 59.8 ± 1.57 . The means, 327.0 and 313.5 ± 4.81 show some difference, but within probability limits. The means 126.0 and 128.4 ± 3.13 are within the limits of chance. The differences in the sigmas, 90.5 and 91.8 ± 2.40 ; and 58.1 and 59.8 ± 1.57 , may be ascribed to chance. If the assumption is made that both tests measure the same thing with approximately the same precision, the inference can be drawn that the samples, showing only chance variability, are adequately representative samples. However, this has not been demonstrated. The coincidence in the

TABLE 2
DIAGNOSIS AND PREDICTION

	% iles*			%**	%**	%	P***	Q ***	
Score	T	D	ND	D	ND	$\sigma \pm$	D	ND	Categories
575-599	99	99	99	100.0	0		.99009	.00991	Clinical delinquents (Pre-clinical delinquents)
500-574	99	99		100.0	0		.99009	.00991	
525-549	99	98		100.0	0		.99009	.00991	
500-524	98	97		100.0	0		.99009	.00991	
475-499	97	95		100.0	0		.99009	.00991	
450-474	96	93		100.0	0		.99009	.00991	
425-449	94	87		100.0	0		.99009	.00991	
400-424	92	83		100.0	0		.99009	.00991	
375-399	87	76		100.0	0		.99009	.00991	
Plimsoll Mark II									
350-374	83	66	99	97.2	2.8	2.74	.97222	.02778	Potential delinquents
325-349	77	57	99	97.5	2.5	2.47	.97500	.02500	(Pre-clinical delinquents)
300-324	72	47	99	97.8	2.2	2.20	.97777	.02223	De facto delinquent
275-299	67	38	97	82.4	17.6	6.54	.82352	.17648	Court delinquent I
250-274	62	29	96	77.4	22.6	7.59	.77419	.22581	Court delinquent II
225-249	56	19	92	73.5	26.5	6.30	.73469	.26531	Proto-delinquent
200-224	49	10	87	64.2	35.8	6.59	.64151	.35849	Prope-delinquent
175-199	43	5	80	42.9	57.1	7.63	.42875	.57143	Chance delinquent
150-174	36	2	69	23.5	76.5	5.94	.23529	.76471	Problem boys
125-149	28	1	56	12.7	87.3	4.49	.12727	.87273	(Schizoids)
Plimsoll Mark I									
100-124	19		37	0	100.0		.00991	.99099	Non-delinquents
75- 99	9		18	0	100.0		.00991	.99099	(Schizoids)
50- 74	4		5	0	100.0		.00991	.99099	
25- 49	2		1	0	100.0		.00991	.99099	
0- 24	1		1	0	100.0		.00991	.99099	
Test Data									
	D						ND	T	
Cases (N)	365						365	730	
Means (M)	315.5 \pm 4.81						128.4 \pm 3.13	220.96 \pm 4.47	
Sigma (σ)	91.8 \pm 2.40						59.8 \pm 1.57	120.72 \pm 2.23	
Validities									
Critical ratio									
(CR)	32.3								
rbis.	.961 \pm .012								
rbis.	.560 \pm .050							(Wide-spread distribution rbis.)	
Reliabilities									
				r			r corr.		
Even-odd (N730)				.932 \pm .005			.965 \pm .003		
Even-odd (N200:									
100 D; 100 ND)				.924 \pm .010			.961 \pm .005		
Test-retest (N200:									
100 D, 100 ND)				.927 \pm .010					
Even-odd (N365 D)				.861 \pm .013			.925 \pm .008		
Even-odd (N365 ND)				.822 \pm .017			.902 \pm .010		

*Percentiles refer to lowest step in interval.

**Actual per cents from finite data.

***Theoretical probabilities.

statistical values could have occurred from the use of an entirely unrelated instrument measuring entirely unrelated traits. It is not probable, however, in this case.

2. *Validities*

Various methods were used for determining these. The critical ratio (*CR*) for *D* and *ND* groups was 32.3, indicating a dichotomy quite clearly. Calculation of the bi-serial coefficient of correlation yielded $.961 \pm .012$. Both measures indicate a degree of test validity. Since the items themselves were selected because they distinguished between *D* and *ND* groups, the inference may be drawn that in some instances item validity is test validity. The test validity in this case is the ability to distinguish a dichotomy. The bi-serial coefficient of $.961 \pm .012$, though an indication of validity, does not show the great amount of overlapping present, as is manifest by inspection of Table 2. A bi-serial coefficient of correlation for wide-spread distribution (5, pp. 384-391), was found to be $.560 \pm .050$, indicating considerable overlapping. To obtain a coefficient of $.961 \pm .012$ by the last method, almost no overlapping could be permitted. Wide-spread bi-serial coefficients of this magnitude are much more rare than equal magnitudes obtained by the first correlation method.

3. *Reliabilities*

These were obtained by various methods. Even-odd reliability, $N = 730$, was $.932 \pm .005$; corrected by the Spearman-Brown prediction formula, it became $.965 \pm .003$. Even-odd reliability ($N = 365$ *D*) was $.861 \pm .013$, corrected to $.925 \pm .008$. Even-odd reliability ($N = 365$ *ND*) equaled $.822 \pm .017$, corrected to $.902 \pm .010$. The greater reliability given by the total group is to be expected since the range is extended. There is little difference in even-odd reliabilities for *D* and *ND* groups. Even-odd reliability ($N = 200, 100, D, 100 ND$) equaled $.924 \pm .010$, corrected to $.961 \pm .005$. Test-retest reliability for the same group equaled $.927 \pm .010$. There is some difference between even-odd reliability and test-retest reliability, when the former is corrected. Test-retest reliability may not be greater than it is because taking the test a second time, in many cases resulting in a test-retest reliability higher than the odd-even, was in this case accompanied by considerable apprehension and sometimes by outright refusal. The test-retest method resembles another method better known by many of the subjects. It may be that the test-retest techniques yielded lower reliabilities for this reason. Inspection of the scattergram for the test-retest method seemed to confirm this. Non-delinquent scores were less variable than scores

for delinquent boys. If calculated, test-retest coefficients for non-delinquents probably would have been higher than for non-delinquents though it was expected that delinquent boys would have made more variable scores since, as scores increase in magnitude, the probability of identity decreases.

On the whole, the reliabilities are adequate. The effort to throw the standard errors out of the hundredths into the thousandths place was not particularly successful.

4. Percentiles

These were determined for *D*, *ND*, and the combined groups. The combined group percentiles lose much of their value since the combined groups do not really represent the total general population; one-half the general population is not delinquent, only one per cent may be expected to become delinquent. In Table 2, the percentiles represent the lowest point in any interval. Table 3 shows the percentiles for *D*, *ND*, and the combined groups to the nearest five points of any score. In many cases this will make little or no difference in the percentiles found. A table showing the percentile for any score was erected but proved too elaborate for publication. It is to be noted that percentiles 99 and 1 have much larger errors in all cases than percentiles clustering about 50; the former errors are about three times as great.

5. Plimsolls

In Table 2, Plimsoll 1 indicates the level to which an individual may be loaded with delinquent traits before he begins to fall in danger of becoming delinquent. No delinquents, with an exception discussed below, was found below the first plimsoll. Plimsoll II is the level above which the individual becomes so loaded with delinquent traits that delinquency becomes almost, but not quite, inevitable. Columns headed by %*D* and %*ND* show the actual number of cases which were found to be delinquent or non-delinquent in any given interval. Above Plimsoll II were no non-delinquents; below Plimsoll I, were no delinquents. This is true only for this set of finite data. It is not possible to say, however, that delinquents will never appear below the first plimsoll or non-delinquents above the second. For this reason, Columns *P* and *Q* were calculated by the La Place theorem

$$\left(P = \frac{n}{n+1} \right),$$

($Q = 1 - P$) to indicate that delinquents can appear below Plimsoll I and non-delinquents above Plimsoll II.

Between the plimsolls, the probability of delinquency varies according to the level. At the interval 350-374, 97.2 per cent were delinquent; and 2.8 per cent, non-delinquent, with an error of 2.74, indicating that the difference in percentage is not due to chance (*vide* column $\% \sigma$, Table 2).

TABLE 3
PERCENTILES (P)

Score	Total P	$\sigma_p \pm$	D-P	$\sigma_p \pm$	ND	$\sigma_p \pm$
600	99	16.68	99	17.42		
595	99	16.68	99	17.42		
590	99	16.68	99	17.42		
585	99	16.68	99	17.42		
580	99	16.68	99	17.42		
575	99	16.68	99	17.42		
565	99	16.68	99	17.42		
560	99	16.68	99	17.42		
555	99	16.68	99	17.42		
550	99	16.68	99	17.42		
545	99	16.68	99	17.42		
540	99	16.68	99	17.42		
535	99	16.68	99	17.42		
530	99	16.68	99	17.42		
525	99	16.68	98	13.89		
520	99	16.68	98	13.89		
515	99	16.68	98	13.89		
510	99	16.68	98	13.89		
505	99	16.68	97	12.05		
500	98	12.91	97	12.05		
495	98	12.91	97	12.05		
490	98	12.91	96	10.93		
485	98	12.91	96	10.93		
480	98	12.91	96	10.93		
475	97	11.20	95	10.15		
470	97	11.20	95	10.15		
465	97	11.20	95	10.15		
460	97	11.20	94	9.58		
455	97	11.20	94	9.58		
450	96	10.16	93	9.13		
445	96	10.16	92	8.77		
440	96	10.16	91	8.47		
435	95	9.44	90	8.21		
430	95	9.44	89	8.00		
425	94	8.89	88	7.81		
420	94	8.89	87	7.64		
415	94	8.89	86	7.49		
410	93	8.49	85	7.36		
405	93	8.49	84	7.24		
400	92	8.49	83	7.13		
395	91	7.87	82	7.04		
390	90	7.63	81	6.95		
385	90	7.63	79	6.79		
380	89	7.40	77	6.66		
375	88	7.26	76	6.60		

TABLE 3 (continued)

Score	Total <i>P</i>	$\sigma_p \pm$	<i>D-P</i>	$\sigma_p \pm$	<i>ND</i>	$\sigma_p \pm$
Plimsoll II						
370	87	7.10	75	6.55	99	12.01
365	86	6.97	72	6.41	99	12.01
360	85	6.84	70	6.33	99	12.01
355	84	6.74	68	6.28	99	12.01
350	83	6.64	66	6.21	99	12.01
345	82	6.54	65	6.19	99	12.01
340	81	6.46	63	6.14	99	12.01
335	80	6.34	61	6.11	99	12.01
330	79	6.31	59	6.08	99	12.01
325	78	6.25	57	6.06	99	12.01
320	76	6.14	55	6.04	99	12.01
315	75	6.09	53	6.03	99	12.01
310	74	6.04	50	6.02	99	12.01
305	73	6.00	49	6.02	99	12.01
300	72	5.96	47	6.03	99	12.01
295	71	5.92	45	6.04	99	12.01
290	70	5.89	43	6.05	99	12.01
285	69	5.86	41	6.08	99	12.01
280	68	5.83	40	6.09	99	12.01
275	67	5.80	38	6.13	98	9.06
270	66	5.77	36	6.17	98	9.06
265	65	5.75	34	6.21	98	9.06
260	64	5.73	32	6.28	97	7.69
255	63	5.73	31	6.38	97	7.69
250	62	5.70	29	6.37	96	7.12
245	61	5.68	27	6.45	96	7.12
240	60	5.67	25	6.55	95	6.62
235	59	5.65	23	6.66	94	6.24
230	57	5.63	22	6.72	93	5.95
225	56	5.62	19	6.95	92	5.71
220	55	5.61	18	7.04	91	5.52
215	54	5.61	16	7.24	90	5.53
210	52	5.60	14	7.49	89	5.21
205	51	5.60	12	7.81	88	5.07
200	50	5.60	10	8.21	87	4.98
195	49	5.60	9	8.47	85	4.80
190	47	5.61	7	9.13	84	4.72
185	46	5.61	7	9.13	83	4.65
180	45	5.61	6	9.58	81	4.53
175	43	5.63	5	10.15	80	4.47
170	42	5.64	4	10.93	77	4.34
165	41	5.65	3	12.05	75	4.27
160	40	5.67	2	13.89	73	4.18
155	38	5.70	2	13.89	71	4.15
150	36	5.73	2	13.89	70	4.13
145	35	5.75	1	17.42	66	4.05
140	33	5.80	1	17.42	63	4.00
135	32	5.83	1	17.42	61	3.98
130	30	5.89	1	17.42	59	3.96
125	28	5.96	1	17.42	56	3.94

TABLE 3 (continued)

Score	Total P	$\sigma_p \pm$	D-P	$\sigma_p \pm$	ND	$\sigma_p \pm$
Plimsoll I						
120	27	6.00			52	3.93
115	26	6.04			48	3.93
110	24	6.14			44	3.94
105	21	6.31			41	3.96
100	20	6.34			38	3.99
95	18	6.54			35	4.03
90	15	6.84			30	4.13
85	13	7.10			26	4.23
80	11	7.40			18	4.59
75	9	7.87			14	4.88
70	8	8.15			10	5.35
65	7	8.49			8	5.71
60	6	8.89			6	6.24
55	5	9.44			5	6.62
50	4	10.16			4	7.12
45	3	11.20			3	7.69
40	3	11.20			2	9.06
35	2	12.92			2	9.06
30	2	12.92			1	12.01
25	2	12.92			1	12.01
20	2	12.92			1	12.01
15	2	12.92			1	12.01
10	1	16.68			1	12.01
5	1	16.68			1	12.01

At the interval 57.1 non-delinquent, with an error of ± 7.63 , indicating that at this level, there is as much chance of becoming delinquent as non-delinquent. The assumption is made that the scores fall at the mid-point of the interval. Actually, of course, they fall within that range.

6. Categories

Table 2 shows three categories: Clinical delinquents, potential delinquents are those found in schools for delinquent boys; the fact that a boy is placed in such a school *ipse facto* makes him a clinical delinquent. The test is really valid for distinguishing non-delinquents from other types of delinquents only insofar as the latter resemble full clinical delinquents. Between the two categories are potential delinquents who fall into sub-categories. The test can distinguish between the sub-categories only partially.

Sub-categories of potential delinquents are: pre-clinical, de facto, Court I, Court II, proto-, prope-, chance, schizoid delinquents, and problem boys. Pre-clinical delinquents are those who score above the second plimsoll mark but are not (yet) in a school for delinquent boys; for a boy to be classified

as a pre-clinical, he must also be classifiable as a proto-delinquent, a sub-category described below; this sub-class is necessary to distinguish those who are in delinquent schools from those who are not, but roam the streets although both have the same characteristics. *De facto* delinquents are former clinical delinquents. Court I delinquents are those who have been brought to court once; Court II delinquents are those who have been brought to court more than once. Proto-delinquents are those who have committed legally punishable acts but have never been apprehended. Prope-delinquents possess many of the qualities of delinquents but who have not committed punishable acts, so far as the test discerns. Prope-delinquents are often neglected children, one step away from proto-delinquents, who are one step away from court, pre-clinical or clinical delinquents. Chance delinquents are those who become delinquent by accidents; they happen to be caught with the goats; they usually score low. Schizoid delinquents also score low, below the first plimsoll mark; they are dreamy individuals with no particular direction or life-plan; they are apt to commit seemingly pointless crimes. Delinquency in schizoids is a secondary problem. Problem boys are those in open conflict with society, the home, the school or all; over 90 per cent of these will become delinquent by the 21st year. Apparently non-delinquent boys who score high are prope-delinquents; their chances of becoming delinquent increase as the score increases.

The test itself can distinguish mechanically only between pre-clinical delinquents, Court I, Court II, proto-delinquents and prope-delinquents. Clinical, *de facto*, problem boys, schizoids, are distinguished clinically. Schizoids are distinguished by low scores and by the schizoid syndrome. The "delinquency of chance" is a post-mortem diagnosis. It is not possible to relate score magnitude to sub-categories.

A fundamental weakness of the norms is the consideration of wide age distributions as one category. Behavior, or test level, which would indicate delinquency in a nine-year-old, may not be as serious as in a 16-year-old and may be less serious in a 21-year-old, i.e., it has less and less meaning for prediction of a delinquent carrier. The item "Do you skip school often?" is a case in point. By adulthood, nearly all have committed at least minor delinquencies. Indeed, some of the characteristics may be measures of maturity. For this reason norms could be greatly improved were they to be determined for each age level. Again, the norms are for white boys. It is strongly suspected that negro boys who become clinical delinquents will have, on the average, lower delinquency scores. Various factors will determine delinquency. The onset of World War II brought an increase; with the

end of the War, came a decrease; the Korean campaign again was accompanied by an increase. A new law may bring apparent increase in delinquency; an old law more strictly enforced may do the same. The effectiveness of a police force is related to bringing delinquency to light. A highly delinquent boy from a responsible home may never become a clinical delinquent; a comparatively non-delinquent boy with no home may be sent to a school because there is no place else to send him. There is nothing especially unjust in this. The point is not to commit a boy, either in the former case or the latter. Commitment may mean crystallization of delinquent traits and attitude. Delinquent society has particularly effective modes of transmitting its culture.

It must also be pointed out when the test has been given, the diagnosis and prediction made, very little has been accomplished; the real task, that of re-adjustment and re-education remains.

D. INTERPRETATION

The personality of the individual may be reconstructed from the test record sheet. A number of such reconstructions are given below as illustrations.

Case I. Age: 14 years, 1 month. Delinquency score (*D*-score): 582.
D % ile: 99. Truth score (*T*): 76 Category: clinical delinquent.

He describes himself as nervous. Losing a game upsets him. Headaches are bothersome. Stomach upset. Things sometimes seem unreal. Muscles twitch. Cannot control feelings. Loses temper easily. Feelings easily hurt. Curses, swears, uses dirty talk. He feels his body is not as perfect as he would like. His friends have turned against him. He cannot get along with his playmates. He is usually restless. Jerks his head and neck. Talks in his sleep. His hands and feet are awkward. He is not good at helping. People are unreasonable. They often neglect him. He is tired at the end of the day but he cannot fall asleep. He feels that others are better looking. He tires of people easily. His mother is always worrying that he'll get sick. His parents are separated. He is usually angry. Knows girls who have been in court. Reports stomach pains and bad dreams. He feels that he is not getting a square deal in life. People are against him. His friends cheat. Nightmares are troublesome. His eyes hurt. He has many colds. His arms are tense. He feels that policemen are watching to get something on him. He has feelings of smothering. Others are unfair to him. He does not like school. Cries easily. People think he is queer. He is happier alone. Has had a steady girl for three months or more. Goes with girls generally. Stays out all night. Reports lots of bad luck. His muscles tremble. Ears buzz. Has been sick from drink. His hands and feet fall asleep often. He gets tired easily. Feels pressures

in head. Is afraid of going insane. Does not fit at school. People don't treat him fairly. He has many problems. People at home are quarrelsome. He tires of work. People find fault with him. Teachers are too strict. People don't like him generally. He often wakes up at night. His hair is wrong. He has been told he can't be good. He wants to get even with someone. Has toothaches. People say he's not nice. His teachers are unfair. His teeth are unattractive. His mouth and lips are wrong. He often stares into space. Asks that questions be repeated. Associates with bad boys. Had fainting spells. His friends drink. People he knows are mean. He calls them names. He has been punished too often. He has many problems. He cannot sit still. Life is not worth it. His friends have been in court. He smokes. Gets upset when he can't smoke. He can't sleep because his heart pounds. He has run away from home. Accused of masturbation. Has had heterosexual experience. His parents get upset because he drinks. He throws up often.

He has been brought to court several times. Has stolen bicycles, held up people; cashed checks not his own; stolen money; stolen from stores; stolen cars; broken into buildings and robbed.

His records check with the delinquencies reported. The truth score, 76 per cent, shows he is quite willing to reveal himself generally. His actions and appearance are that of a "hard guy."

Case II. Age: 12 years, 1 month. *D* score: 504. *D* % ile: 97. *T* score: 82. Category: clinical delinquent.

He has spells of sneezing. Reports himself as nervous. Has gas on stomach. Things sometimes seem unreal. Parents are separated. He is not usually cheerful. Easily discouraged. He feels uneasy on a bridge. Has stayed out all night. Friends have turned on him. He has been away from home. His parents expect too much of him. People are generally unreasonable. They neglect him. He is ashamed of his father's work. Others are better-looking than he is. People say they cannot depend on him. He is usually angry. Skips school a lot. His girl friends have been in court. His friends cheat. Nightmares are troublesome. Feelings of smothering reported. Day-dreams are usually about unpleasant things. He has shooting pains in his head. Has many sore throats. Has had a steady girl. Goes with girls generally. His muscles tremble. He catches cold easily. People say he lies. He gambles, shoots crap. Has hurt someone with a knife. Has been sick from drink. He gets tired easily. His hands fall asleep. He vomits frequently. Is afraid of going insane. Does not fit at school. People treat him badly. Those at home are quarrelsome. He is unusually unlucky. His teachers are too strict. People don't treat him well generally. He has to start a fuss to get what's coming to him. His teachers are unfair to him. He is making poor progress in school. His face is unattractive. His mouth looks wrong. His companions are bad. They drink. He calls people names. His teeth are ugly. He wets his bed.

He doesn't like people generally. He has stayed out all night. Sometimes his head aches badly. He feels that he has many problems. Life often is not worth the trouble. His friends have been arrested. He smokes. Gets upset when he can't smoke. His parents get upset because he drinks. He has been in court repeatedly. Throwing up is a problem.

He has stolen things generally; stolen from stores; stolen money, bicycles, cars; cashed checks not belonging to him; has hurt someone with a knife; has broken into buildings.

His record confirms the delinquencies.

Case III. Age: 12 years, 11 months. *D* score: 246. *ND %* ile: 96. *D %* ile: 27. *T* score: 78. Probability of delinquency: 73.5 in 100. Category: Potential delinquent. Sub-category: Proto-delinquent.

He says others seem healthier than he. It upsets him to lose. He is easily discouraged. Loses his temper. Is afraid of hell. Uneasy on bridge. His feelings are easily hurt. He feels his body is not as good as he would like. He talks in his sleep. Reports head tics. People are unreasonable. They neglect him. People say he's undependable. His friends cheat. He has nightmares. He thinks Sunday School teachers are odd. Others say he is unfair. He acts on impulse. He has had bad luck. People say he says things that are not true. He shoots crap and gambles. He is afraid of insanity. People at home are quarrelsome. People generally do not like him. He does not believe in doing things for others. He wants to get even with several people. People generally have mistreated him. He has to fight to get what's coming to him. He has been accused of not being nice. He is not making good school progress. His companions are bad. He calls people names. He gets tired of his friends easily. He has been punished too often. He finds it hard to sit still. Life is not worth it. He smokes. He can't sleep because his heart pounds. His mother is very hurt by what he does.

He has stolen things from stores; has stolen money and bicycles. The truth score shows that he is fairly willing to reveal himself. His chances of becoming clinically delinquent are 73.5 in 100; his chances of remaining a non-clinical delinquent are 26.5 in 100. Were he not white, his chances of becoming a clinical delinquent, at his *D* score level, would have been greater, to an undetermined degree.

Case IV. Age: 12 years, 5 months. *D* score: 165. *ND %* ile: 75. *D %* ile: 3. *T* score: 52. *P*: 23.5 in 100. Category: potential delinquent. Sub-category: Prope-delinquent.

He has been ill quite a bit as a child. His friends have turned against him. People expect too much of him. People are unreasonable. He has stomach pains. Has the habit of drumming with his fingers. He can't sleep. His arms are tense. He feels that policemen watch boys. He is often lost in thought. He acts on impulse. He catches cold easily. Complains of buzzing in his ears. People say he lies. His

mother worries about him. His feet fall asleep. He does not fit at school. He is unusually unlucky. People find fault with him. He does not believe in doing things for others. His teeth ache. People don't treat him well. He is accused of not being nice. He associates with bad boys. He calls people names. His legs are tense much of the time. He finds it hard to sit still. His face is unattractive. Life often is not worth living. His mother is very hurt by what he does.

The *D*-score shows that his chances of becoming delinquent are 23.5 in 100. However, the *F* score of 52 indicates that he is revealing himself only about 50 per cent of the time. He is hiding unflattering aspects of his personality. His real score is some place above 165, at an undetermined point.

Case V. Age: 12 years. *D* score: 48. *ND* % ile: 3. *T* score: 84. Category: non-delinquent.

He reports himself as somewhat nervous. He helps people at having a good time. Talks to strangers. Has hay-fever. His feelings are easily hurt. He has nightmares occasionally. He complains of too much work. He has a few close friends. He feels his face is somewhat unattractive.

The *T*-score shows that considerable faith may be placed in the *D*-score. He is willing to reveal himself by admissions uncomplimentary to himself, yet shows very few delinquent characteristics. The score is confirmed by personal impressions.

REFERENCES

1. BODIN, M. Do problem children become delinquent and criminal? *J. Crim. Law & Criminol.*, 1936, **27**, 345-559.
2. CRONBACH, L. J. On estimates of test reliability. *J. Educ. Psychol.*, 1943, **34**, 485-494.
3. DALE, E., & CHOLL, J. S. A formula for predicting readability: Instructions. Educational Research Bulletin, College of Education, Ohio State University, 1948.
4. KELLEY, T. L. The Kelley Statistical Tables. New York: American Book, 1935.
5. PETERS, C. C., & VAN VOORKIS, W. R. Statistical Procedures and Their Mathematical Bases. New York: McGraw-Hill, 1940.
6. ZAKOLSKI, F. C. Studies in delinquency: II. Personality structure of delinquent boys. *J. Genet. Psychol.*, 1949, **74**, 109-117.
7. ————Studies in delinquency: II. Prediction of delinquency in boys. *J. Genet. Psychol.*, 1949, **74**, 119-123.
8. ZUBIN, J. Nomograph for the determining of the significance of the difference between the frequencies of events in two contrasted series or groups. *J. Amer. Stat. Assos.*, 1939, **34**, 359-344.

Teachers College of Connecticut
New Britain, Connecticut

COMPENSATORY BEHAVIOR AS COMPLETION OF THE SENSORY-MOTOR CYCLE: AN INTERPRE- TATIVE HYPOTHESIS*

Department of Psychology, Long Island University

HENRY GUZE

The problem of compensatory behavior has significance for the interpretation of a large array of psychological phenomena. Some authors, Adler (1) for example, have built up an interpretative psychological system in terms of responses apparently made to compensate for physical or behavioral inadequacies.¹ At times the problem presents itself as to how it might be possible to explain compensatory behavior in mechanistic terms which place a minimum of emphasis upon the usual purposivistic interpretations. This paper concerns itself with the concept of compensation as completion of an incomplete sensory-motor cycle. If this cycle cannot be completed in terms of motor expression, it may be that the sensory aspect is perpetuated in terms of "feed-back"² as implied by Ashby (2). Such feed-back takes the form of a circuit much along the lines of the Watt engine (2). Thus *A* affects *B* which affects *C* which in turn affects *D* which finally affects *A*. Lorente de No (21) has emphasized this and demonstrated histologically the widespread occurrence of neuron circuits. Ashby (2) quotes Lorente de No's "Law of Reciprocity of Connections" thusly: "If a cell complex *A* sends fibres to cell complex *B*, then *B* also sends fibres to *A*, either directly or by means of one internuncial neuron." It is the opinion of this writer that a circuit in the organism's brain would be brought into action particularly if the stimulus, externally or internally induced, could not be "released" in terms of a neuro-motor response.³ From this standpoint something can

*Received in the Editorial Office on November 11, 1951.

¹It is to be remembered that Adler's *Study of Organ Inferiority and Its Psychical Compensation* (1) is based upon biomedical data that begins with a genetic point of view. The current paper is not concerned with the same problem but rather with the mechanisms of compensatory action. However, since such action occurs when there is something interfering with the usual reaction pattern, interference could of course, be due to an organ inferiority (genetic or induced) in some cases [see reference to Lashley (18) later in paper].

²Feed-back, as will be shown later, can occur also as a manifestation of the complete sensory-motor cycle where the motor response becomes the stimulus for a new cycle.

³Some circuits, particularly those involving the whole body have sensory-motor components, as will be shown later. Other circuits may be only sensory in nature.

be said about what is meant herein by compensatory response. Such response is that which completes a sensory-motor cycle by the indirect route, that is to say, if an individual's stimulus complex is designed to elicit an arm flexion which he is prevented from giving, he may respond with activity of some other part of the body. If there is prevention of all response, there will be a reverberatory circuit set up which continues actively⁴ until such times as gross response is possible. Such circuit need not of course depend only upon this type of stimulus initiation, and may occur under other circumstances. Among others, Jennings (16) is responsible for the idea that movement of the organism brings about stimulation which in turn causes further movement, etc. Thus the statement of Ashby (2) "The living organism and its environment together form a system containing feed-back," is clarified. Such complete sensory-motor circuit might theoretically account for continuous bodily action in behavioral terms. Inhibition of the motor aspect of this cycle in terms of skeletal or somatic response might lead, on a hypothetical basis, to either a visceral⁵ cycle being formed, or a compulsive drive for action appearing (or both). This action drive need not have ideational content, since there is no reason to assume that a reverberatory circuit is ideational (although it may be). The ideational content may become attached experientially in terms of previous release of the tension, perhaps a conditioning factor through temporal continuity.

This would be true of those which operate in areas of the cortex. In such case each motor component could exercise, likewise, a sensory action. While this is true of the whole body response, as indicated above, the "motor" response in the cortex is merely so named because it is discharged by another neuron. Strictly speaking it is not a motor response.

⁴Ashby (2) says that a system to be "goal-seeking" must possess feed-back.

⁵This would be psycho-visceral in that the sensory aspect would bring about visceral motor reaction through the autonomic efferent. This in turn might cause further sensory stimulation and initiate a full cycle. The psycho-visceral response is a compensation for inability to respond in terms of the somatic component when such response is appropriate. It is probable that this may be a means of interpreting the autonomic and visceral changes in the fear or anger situation as described by Cannon (7). There is evidence to warrant the statement, on empirical grounds, that when the organism is able to respond with gross bodily action to an intense stimulus, the reaction of fear or rage does not occur. It would appear to this writer that the distinction between fear and rage is sheerly a quantitative one which takes on qualitative meaning (albeit often the two may be confused). The fear response is more intense than the rage response in that, fear results from an inadequate repertoire for skeletal response to a given stimulus complex whereas rage results from an inhibition of a response to which there may be resistance within the individual due to previous training or resistance from the environment due to physical barrier. The latter aspect does not always result in obvious rage. Contrary to the frustration-aggression concept (8), no overt hostility need appear. It is necessary to distinguish between rage and rage with expression in terms of hostility. Autonomic nervous differences in fear and rage may be due to threshold differences in the autonomic system.

What is a compensatory response? It is (as used in this paper) a response made by one part of the body when another part which is being stimulated cannot for social or personal reasons make an appropriate response. For example a subject retaining urine in the full bladder, moves the feet or legs. This is compensatory in that there is thus some aspect of completion of the sensory-motor cycle, albeit through an inappropriate channel. Here the drive (sensation) persists, although it is momentarily perhaps reduced. In this case, of course, the compensatory response will not replace the appropriate response because there is no real reduction of drive tension or sensory stimulation. On the other hand a response to another sensation, albeit inappropriate from the standpoint of society or efficiency may become typical because it reduces completely, or very much so, the stimulus or sensory tension.⁶ Thus it is quite clear that such compensatory response as auto-erotism may persist because it successfully reduces the tension of a sensory experience. Thus, if this has become the usual mode of response, it may be difficult to replace it unless another mode of response is introduced which is actually easier to perform. This, of course, requires introduction of an alternative motor response which is designated as appropriate either because of physiological adequacy or social acceptance. Evidence from everyday observation seems to indicate that many compensatory actions occur both in minimal muscle response and gross body movement. These actions may be normally expressed in that they are representative of responses of certain large numbers of organisms inhibited in a given way in a given population. Such normalcy may be pathological in that it is injurious to the biology of the organism, or non-pathological in that it is not injurious. It is necessary to recognize from this standpoint that deviant or "abnormal" behavior may be non-pathological (11).

That the compensatory response may become habit in an organism is well known, and that this may replace and supersede the response which is biologically (without any training of any kind) given, is exemplified in everyday life. It need not be assumed that the compensatory response (as indicated above) need be re-channelized to the basic⁷ response in all cases, since the

⁶Hull (15) has pointed out that the state of an organism's needs plays an important rôle in the causal determination of which of the many habits possessed by an organism shall function at a given moment. Those which are most frequently need-reducing would, of course, be reinforced. Should they be blocked, the next most need-reducing habit might come into play, etc.

⁷The basic response is that which is presumably genetic. Contrary to Ashby (2) and others, this writer believes that this response is not always self-preservative for the organism, as in homeostasis (7). For example, the basic response of a moth (22) is to fly toward a fire under certain circumstances. A compensatory

conditioning and learning situation involves much establishment of compensatory responses which in time come to take on the full tension-reduction value of basic responses to basic stimuli (when these are tension reducing).

Various types of phenomena in the organism are accompanied by spontaneous responses which might be termed "compensatory." Schneck (26) has demonstrated that subjects give spontaneous responses to hypnotic induction. These seem to be motor completions of various bodily sensations. Such motor completion would appear to be involved in many psychodynamic mechanisms⁸ and can act to start a circuit, since the response in itself is sensory to the organism at the same time that it is motor.

Phylogenetically speaking, it would appear that compensatory motor activities can be shown to exist throughout the animal kingdom. If a particular response is blocked, another part of the body will respond to the stimulus or the whole body will respond.⁹ In the lowly protozoan the compensatory

motor response to this light sensation would indeed be desirable to maintain preservation of the organism.

⁸For example projection involves, from this standpoint, a compensatory response in that there is a resistance to the response actually elicited. The sensory-motor cycle is thus completed by a motor response that may distort the sensory interpretation. This does not deny sensory-perceptual distortion as a function of other factors, although the disturbed percept seems to this writer a compensatory motor completion of a sensory motor cycle. This view does not contradict the sensory-tonic field theory of Werner and Wapner (29).

⁹Conditioning research seems to give supporting evidence to this point of view. Harlow and Bromer (12) demonstrated that in monkeys whose cortex was partially "frozen" by drug usage, conditioned responses occurred after the drug wore off, but not in the limb which was being conditioned. The responses occurred in the eyelids, head, and neck. Morgan (23), in commenting on this says: "Thus it was clear that learning had occurred, although the complete motor system from the cerebral cortex on down was paralyzed." The fact that the "learning" was as removed, as it was, from the source of stimulation suggests to this writer that the phenomenon was an aftermath of compensatory overflow. This problem may be behind the audiogenic seizure (24). This phenomenon suggests that, under certain circumstances, there are threshold changes which make spread of neural effect through the nervous system possible. It is noteworthy that a 24-hour food or water deprivation increases reactivity to audiogenic seizure. Is this due to a summation effect because of an increase in activity from hunger alone plus the need for response to the audiogenic factor?

Probably response generalization is also a compensatory phenomenon. As pointed out by Hilgard and Marquis (13) "... A stimulus which has come through training to elicit a particular response may, under some circumstances, elicit a different response without special training. It is evident that generalization is important in the adaptive economy of the organism because environmental situations never recur in nature without change." It was demonstrated by Bechterev (3) that a dog whose conditioned limb is fastened down will lift another limb. This is more than mere equivalence. It is clearly compensatory. Lashley (18) has shown this in a monkey who solved a manipulation problem with the non-trained hand after paralysis of the hand used in this problem. This also appears to be compensatory. The neurology underlying this is not entirely clear but is nonetheless understandable to the extent that mass action and vicarious "function" would indicate that to some

behavior may be a pseudopod formation in the amoeba, or a change of movement in the paramecium. It is appropriate to refer to such behavior as compensatory because the activities in response to block or inhibition are means of "making up" for the inhibition in terms of another reaction.

Compensatory response, as here described, may in fact be facilitative in the learning situation under various circumstances. For example, in an old experiment, Triplett (28) introduced a partition of glass between minnows and larger fishes (perches) which ordinarily were fed upon minnows. After bumping into the glass, the large fishes swam in the opposite direction (or away from the minnows.) They received the sensation of the bump again and again following the motor response to the visual stimulus of the minnows. The change in body movement is a completion of the sensory-motor cycle and is compensatory in that the original response cannot be made. This situation seems to favor the establishment of conditioning which becomes a compensatory response in place of the basic unconditioned response. It might be possible theoretically, for this reason, that extinction occurs when the compensatory response to the original stimulus is no longer necessary. This would apply in a conditioning with reward as in the case of Liddell's (19) sheep. It might also be possible to extend it to cover the general picture of classical conditioning. Under such circumstances the *CS* acts as a signal for the eating response (where food is used.) Since this cannot be made, the salivation alone, which is the compensatory response, appears in its place. Ultimately through lack of reinforcement the sensation of the sound of the bell¹⁰ becomes more localized in its stimulus character and no widespread response is necessary. This is so because the bell is responded to only in terms of the local auditory action and its accompanying occurrences. Previously, through temporal contiguity, one of the responses involved was eating. When the eating did not occur, the compensatory response (for eating) of salivation following the bell.

In this relationship, the theory of repression becomes very important and

extent any part of the organism under central control might take over the action of another part. Thus as shown by Kellogg (17), there is probably a response gradient, which is intensified by inhibition of the usual response. Thus if response *A* cannot be given, *A*₁ which is closest in innervation would appear, then *B*, then *B*₁, etc., on the basis of similarity of function and innervation in a direction of increasing distance from the source of the original response.

¹⁰A stimulus may increase in its intensity value by coupling with another stimulus. When the latter is removed, the former may act (on its own) to stimulate the response which the original *US* elicited (as well as its own response). This response, however, has only part of the strength of the original. From a compensatory viewpoint the *CR* completes a sensory motor cycle which is elicited by the *CS*. In line with the foregoing, Hollingworth's (14) concept of redintegration might be examined.

the view of Estes (9) that punishment does not extinguish a response becomes clear. This is so because as punishment is intensified, the need for motor response is likewise increased. This motor activity may be turned inward in the form of anxiety which is of autonomic visceral origin. The sensation of anxiety in itself stimulates the organism and there is need for response. This may be in terms of generalized activity on the skeletal somatic level, or it may be in terms of visceral discharge, perhaps as hypertension. It must be noted that in these areas of activity it is likely that a skeletal-visceral feed-back system will be established.

The neurotic organism may respond either by random overt activities in completing the sensory-motor cycle or by specific activities which are peculiar and difficult to understand in terms of the neutral observer and of the patient himself in most cases.¹¹ These activities become fixed in the organism as a function of their efficiency in completing the sensory-motor cycle.¹² This concept agrees with the view of Freud (10) which assumes that there is symptom conversion of anxiety. This symptom conversion is response to the stimulation in terms of some motor act which completes the sensory-motor cycle that is set up. If the stimulus that initiates anxiety is not eliminated, then the entire set of events persists. Anxiety can only occur when no motor response to a stimulus is available or permissible, or when those responses made do not fully complete the sensory-motor cycle. In this sense a limited amount of anxiety is characteristic of the healthy organism and is probably a descriptive characteristic of motivated behavior. Perhaps it is unnecessary to translate the concept of desire into terms of sensory-motor completion, however such approach would reduce much teleological thinking in regard to the motivated organism. Desire is a sensory aspect of the sensory-motor cycle. Often the standard psychiatric and psychological literature is loaded with untenable and unnecessary descriptions of phenomena which are easy to interpret in terms of the sensory-motor cycle. These descriptions give an odd kind of independence to psychological or psychosomatic occurrences, presenting them in reified terms that are scientifically difficult to define and accept.

¹¹The psychosomatic case is marked by inhibition of motor response on an overt level or by overstimulation on the sensory level and thus a physical inability to complete the numerous sensations set into action.

¹²The act becomes reinforced by practice and thus becomes the typical and preferred type of response. This means often that a more socially sanctioned response will be completely neglected and resisted. Only in so far as the compensatory response is harmful to the organism or society in a biological sense would this writer consider such condition pathological (11).

Rubin (25), in regard to disorder of the gastro-intestinal tract, says:

In a sense then people who, because of the inner conflicts described, are unable to express themselves outwardly and are forced to repress these overstrong receptive wishes, ultimately bring these wishes to expression in the tacit physiologic language of the stomach functions. The stomach behaves in response to this stimulus as if it were taking or about to take in food. The more strongly the unconscious receptive wishes are repressed, the greater will be the unconscious striving for receiving help and love. Food, then, is wanted not for the satisfaction of hunger, but as a symbol for love and help.

The foregoing description is interesting and useful, but is loaded unnecessarily in terms of teleology. Is it not more simple to explain the gastro-intestinal reaction sheerly as completion of a sensory-motor cycle in which overt bodily motor action is inhibited? Rubin, himself, implies this, only to go on and later lose himself in the language of the Freudians. One might ask two questions in his defense. Why is the gastro-intestinal tract "chosen" for the response, and why is there a need for food? The first question can be hypothetically answered by the assumption that the gastro-intestinal tract became involved in a given case because at the time of the stimulation its threshold for response was lower than that of the other viscera. This lowering of the threshold, for any bodily system or part of a system, may be temporary as a result of exogenous factors, trauma or fatigue, genetic factors or pre-natal factors [see Sontag (27)]. Once the gastro-intestinal response is established to the conflict, it may become fixed by conditioning or some similar phenomenon. The sensation the organism derives from the gastro-intestinal changes is that of anxiety, a sensation which is descriptively similar to the hunger activities of the tract. The subject responding to his bodily discomfort wants food, his typical remedy for gastro-intestinal disturbance of this kind.

When a response cannot be made or cannot easily be made, it is the thesis of this paper that compensatory bodily responses are made. Such responses disappear overtly and become minimal in internal expression as the original act¹³ becomes integrated into the personality as a result of learning unaccompanied by emotional resistances. This concept follows the data in regard to the acquisition of skills as being marked by a decrease of gross bodily action and an increase in specificity of response as well as integration (5, 6). The compensatory response becomes most outstanding when an act cannot be

¹³This refers to the act which elicits the compensatory behavior. Failure to master successfully the original act results in increased compensatory expression.

performed, albeit the stimulation for its performance is intense. Behavior mechanisms, as is well known, are often compensatory (see earlier in paper), and it might be said that the projective techniques depend much upon this aspect of human activity. On the Rorschach, for example, an *F*— may be the means of evading another response which the subject cannot give because of emotional resistance. Distortion appears as a means of responding compensatorily. As time goes on, it may become a mark of the personality and a clue to disorder. It also becomes so habitually reinforced that it may prove most complex to eliminate it. Regression, from this point of view, is compensatory in that the subject responds by means of an earlier repertoire of responses because he cannot, for some reason utilize responses which he has more recently acquired. In this case compensation is on the basis of the older response. This follows Jost's law (30). The newer response cannot be made efficiently. It is thus abandoned and the easier or more stable response appears.¹⁴ It is noteworthy that in neurological disturbances of various kinds and in senility the older responses replace the newer ones which cannot be made. In the psychotic both this "return" to earlier behavioral response as well as the appearance of apparently unrelated compensatory acts marks the picture of social and emotional failure.

Many activities of the organism which have been difficult to understand may be approached or interpreted on the basis of this concept of compensation. The concept of compensatory response can be brought actively into the interpretations of psychosomatic disorder and perhaps even further extended in the interpretation of other disease entities. Neurologically, it would seem indeed that a stimulation of one area of the body may elicit response at some distant focus in defiance of what would seem to be the limits of segmental spread of effect. Livingstone (20) discusses this sensory spread of effect as being currently difficult to interpret. Compensatory response may be a factor, in that blocking of responses in one area will tend to spread the stimulation and elicit responses elsewhere. Thus a diseased organ may lose its usual range of motility and the stimulation of the area may be carried elsewhere for response. The physiological dynamics of the process are well worth further elaboration.

¹⁴As indicated by Woodworth (30) Jost in attempting to explain the advantage of spaced over massed learning found that the older association profited more than a new one from exercise or repetition. While the situation in the text above is not quite the same, the older responses seem to profit more by practice than the newer ones and in regression are reinforced by learning as compensatory behavior.

SUMMARY

It has been hypothesized that compensatory behavior on a physiological or psychological level is the result of blocking of the most usual motor component of a given sensation. The result is a motor component that differs from the original or "natural" one and thus compensates for it. There is a gradient of motor responses of compensatory nature in terms of the various response blockings present, thus a compensatory response often may be far removed from the original or "natural" reaction. Such response may be bizarre as in some pathology or sheerly different but satisfying. The area is open for study in terms of psychosomatic response as compensatory reaction and in terms of referred sensation and response in disease entities. At the behavioral and perceptual level it is a factor in projective distortion and psychopathology.

REFERENCES

1. ADLER, A. Study of Organ Inferiority and Its Psychical Compensation. New York: Nervous & Mental Disease Publishing Company, 1946.
2. ASHBY, W. R. The cerebral mechanisms of intelligent action. In Derek Richter (Ed.), *Perspectives in Neuro-Psychiatry*. London: Lewis, 1950.
3. BECHTEREV, V. M. General Principles of Human Reflexology. New York: International, 1932.
4. BRYAN, W. L., & HARTER, N. Studies in the physiology and psychology of the telegraphic language. *Psychol. Rev.*, 1897, 4, 27-51.
5. ———. Studies on the telegraphic language: The acquisition of the telegraphic language. *Psychol. Rev.*, 1899, 6, 345-375.
6. BOOK, W. F. The Psychology of Skill With Special Reference to Its Acquisition in Typewriting. New York: Gregg, 1925.
7. CANNON, W. B. Bodily Changes in Pain, Hunger, Fear, and Rage. New York: Appleton, 1927.
8. DOLLARD, J., DOOB, L. W., MILLER, N. E., MOWLER, O. H., & SEARS, R. Frustration and Aggression. New Haven: Yale Univ. Press, 1939.
9. ESTES, W. K. An experimental study of punishment. *Psychol. Monog.*, 1944, 57, No. 3.
10. FREUD, S. A General Introduction to Psychoanalysis. New York: Liveright, 1920.
11. GUZE, H. Toward an absolute psychopathology. (Unpublished.)
12. HARLOW, H. F., & BROMER, J. A. Acquisition of new responses during inactivation of the motor, premotor, and somesthetic cortex in the monkey. *J. Gen. Psychol.*, 1942, 26, 299-313.
13. HILGARD, E. R., & MARQUIS, D. G. Conditioning and learning. New York: Appleton-Century, 1940.
14. HOLLINGWORTH, H. L. Abnormal Psychology: Its Concepts and Theories. New York: Ronald, 1930.
15. HULL, C. L. Principles of Behavior. New York: Appleton-Century, 1943.
16. JENNINGS, H. S. The Behavior of the Lower Organisms. New York: Columbia Univ. Press, 1923.

17. KELLOGG, W. N. "Positive" and negative conditioning without contraction of the essential muscles during the period of training. *Psychol. Bull.*, 1939, **36**, 575.
18. LASHLEY, K. S. *Brain Mechanisms and Intelligence*. Chicago: Chicago Univ. Press, 1929.
19. LIDDELL, H. S., JAMES, W. T., & ANDERSON, O. D. The comparative physiology of the conditioned motor reflex: based on experiments with the pig, dog, sheep, goat, and rabbit. *Comp. Psychol. Monog.*, 1934, **11**, No. 51.
20. LIVINGSTON, W. K. *Pain Mechanisms: A Physiologic Interpretation of Causalgia and Its Related States*. New York: Macmillan, 1947.
21. LORENTE DE NO, R. Analysis of the activity of the chains of internuncial neurons. *J. Neurophysiol.*, 1938, **1**, 207-244.
22. MAST, S. O. *Light and the Behavior of Organisms*. New York: Wiley, 1911.
23. MORGAN, C. T., & STELLAR, E. *Physiological Psychology*. New York: McGraw-Hill, 1950.
24. MORGAN, C. T., & GALAMBOS, R. Production of audiogenic seizure by tones of low frequency. *Amer. J. Psychol.*, 1942, **55**, 555-559.
25. RUBIN, S. Gastrointestinal system. In Dunbar, F. (Ed.), *Synopsis of Psychosomatic Diagnosis and Treatment*. St. Louis, Mosby, 1948.
26. SCHNECK, J. M. Psychosomatic reactions to the induction of hypnosis. *Dis. Nerv. System*, 1950, **11**, 1-4.
27. SONTAG, L. W. Determinants of predisposition to psychosomatic dysfunction and disease: Problem of proneness to psychosomatic disorder. In Dunbar, F. (Ed.), *Synopsis of Psychosomatic Diagnosis and Treatment*. S. Louis: Mosby, 1948.
28. WATSON, J. B. *Behavior: An Introduction to Comparative Psychology*. New York: Holt, 1914.
29. WERNER, H., & WAPNER, S. Sensory-tonic field theory of perception. *J. Personal.*, 1949, **18**, 88-108.
30. WOODWORTH, R. S. *Experimental Psychology*. New York: Holt, 1938.

Department of Psychology
Long Island University
380 Pearl Street
Brooklyn 1, New York

SHORT ARTICLES AND NOTES

The Journal of Genetic Psychology, 1953, **83**, 303-306.

THE INTELLECTUAL FUNCTIONING OF A POST- INSTITUTIONAL GROUP*

Brooklyn, New York

ARTHUR D. HAGGERTY

A. INTRODUCTION

This study has as its purpose an evaluation of the level and type of intellectual functioning of the client population of an agency. Clinical judgment and a perusal of the results of psychological tests administered during institutionalization revealed that this group apparently was functioning at a subnormal intellectual level. The results of a small number of Stanford-Binet Scales administered after discharge from institutionalization pointed to the same conclusion. It was then decided to test a group which might be representative of the whole agency population with the Wechsler-Bellevue Intelligence Scale, Form I.

The agency involved in this study is responsible for the after-care of all children and adolescents who are discharged from institutional care (this includes foster homes) in the Diocese of Brooklyn. Such supervision includes vocational, educational, social, and spiritual guidance and planning, and is carried out by a staff of case workers, a clinical psychologist, and psychiatric consultants.

Since intellectual functioning is one of the important factors entering into this planning, a study is being carried out in this area. Studies have already been performed concerning the intellectual functioning level of institutional and post-institutional groups (2, 4, 9), but little has been done concerning an analysis of this functioning. Marsh's study with infants' improvement of *IQ's* after placement in good foster homes adopts the hypothesis that while an institutional group might be functioning at a retarded level, its potential is higher (8). The same hypothesis underlies this present study with regard to a post-institutional group.

In addition, since the deleterious effects of institutionalization are well

*Received in the Editorial Office on November 24, 1951.

known (1, 3, 5, 6), we are accepting the hypothesis that such malfunctioning as exists does so because of emotional disturbance rather than retardation due to any inherent limitation. It is expected that an analysis of the Wechsler results will reveal pertinent data in this direction (10).

B. PROCEDURE

A random group of clients who had been seen for routine psychological examinations was chosen. Since many of the clients referred for testing were considered "problem" children, it may be questioned whether this group is representative of the agency population. The writer, however, believes so since many others were not referred because of any special problems, but because they were available for testing, or because they showed promise, were workable, and as such could definitely profit by psychological study.

The group was composed of 21 males and 13 females between the ages of 11 and 21 years inclusive, with a mean age of 15.2 years. Each individual was administered the Wechsler-Bellevue Intelligence Scale, Form I, and the results were tabulated. Mean scores and standard deviations were computed for each subtest and the three Wechsler *IQ*'s. A mean discrepancy between Verbal and Performance functioning was also computed. The results were compared with the comparable results Wechsler found on his standardization group of 15-year-olds.

C. RESULTS

The results of this study are depicted in Table 1. Here can be compared the results of our investigation group (hereafter called Group *I*) with those of Wechsler's standardization group (hereafter referred to as Group *S*). From an analysis of subtest scores we find that Group *I* is inferior to Group *S* on every subtest except Object Assembly, the subtest upon which children at all levels of intelligence often do well. This also is the only subtest on which Group *I* reached the average level. Similarly, the three mean *IQ*'s for Group *I* are much lower than those for Group *S*, with Verbal and Full Scale functioning falling below the lower limit of average intelligence. The respective standard deviations reveal that Group *I* is more homogeneous and therefore more uniformly inferior on Comprehension, Digit Span, Picture Completion, Block Design, and Performance functioning, while being less homogeneous on Information, Arithmetic, Similarities, Picture Arrangement, Digit Symbol, Verbal functioning, and Full Scale functioning. Thus in these latter areas the investigation group is capable of a wider spread of levels than the standardization group.

TABLE 1
COMPARISON OF SUBTEST SCORES AND *IQ*'s WITH THEIR RESPECTIVE STANDARD DEVIATIONS

	Group <i>I</i>		Group <i>S</i>	
	Mean score	Mean <i>SD</i>	Mean score	Mean <i>SD</i>
Information	6.1	3.85	10.6	2.61
Comprehension	7.7	2.43	9.8	2.93
Digit Span	6.1	2.31	9.0	2.63
Arithmetic	4.2	10.73	8.9	2.54
Similarities	5.8	2.80	9.6	2.13
Vocabulary	5.8	2.80		
Picture Arrangement	8.5	7.98	10.8	3.10
Picture Completion	7.5	2.64	9.6	2.95
Block Design	7.9	2.83	10.7	3.18
Object Assembly	9.8	2.56	9.8	3.45
Digit Symbol	6.6	3.68	10.6	2.42
Verbal <i>IQ</i>	81.8	19.80	100.0	14.60
Performance <i>IQ</i>	92.0	13.34	100.5	15.70
Full Scale <i>IQ</i>	85.0	16.88	100.0	14.57

TABLE 2
COMPARISON OF DISCREPANCIES BETWEEN VERBAL AND PERFORMANCE *IQ*'s

	Group <i>I</i>	Group <i>S</i>
Mean discrepancy between scales	13.4 pts. (Range-2-37)	8.8 points
Discrepancy between mean <i>IQ</i> 's	10.2 pts.	.64 points

It is also apparent, as seen in Table 2, that Group *I* has a much larger discrepancy between Verbal and Performance functioning than Wechsler found in his standardization group, both when considering the mean difference between the scales, and also the difference between the means of the two *IQ*'s.

D. DISCUSSION

There is no doubt that this study points to the inferior functioning of the investigation post-institutional group. It is inferior on most subtests and on all three *IQ*'s to a standardization group of the same mean age, and thus inferior to "normal" criteria. On some subtests and on two of the *IQ*'s Group *I* is more homogeneous and thus uniformly inferior; on others it is much less homogeneous and therefore capable of wider variation.

What perhaps is more important for our discussion is the comparison of average discrepancy between Verbal and Performance functioning. It is apparent that the average individual in Group *I* has a relatively large discrepancy, and is not retarded equally in all subtests or in the two types of intelligence—a condition found more frequently in emotional disturbance than in inherent retardation. Approaching the group's functioning in this

manner resembles that of Jastak who holds that the greatest success an individual achieves (the altitude) is a more valid index of his actual intelligence than the Full Scale *IQ* (7). Thus an evaluation of the subtest altitudes in this case might very well reveal that the group's potential is far above its present functioning level as measured by the Full Scale *IQ*. This would seem to be especially true in light of the extreme variations in subtest and Verbal-Performance functioning levels. Such an approach has important implications when considering therapeutic services for a group of this type, and attempting to raise its intellectual functioning closer to its true potential level.

E. SUMMARY

The Wechsler-Bellevue Intelligence Scale, Form I, was administered to a group of 34 clients of an after-care agency. Subtest scores and *IQ*'s were tabulated and means and standard deviations computed. These results were then compared with comparable data from a Wechsler standardization group. An analysis of the patterns formed reveals that the investigation group does function below normal, but that the extreme variation is more compatible with emotional disturbance than with inherent limitation. Thus an approach utilizing the concept of the altitude rather than the *IQ* as an evaluative measure provides a more optimistic outlook toward such a group, and supplies a framework for remedial therapeutic measures.

REFERENCES

1. CLARDY, E. R., & HILL, B. C. Sleep disorders in institutional disturbed children and delinquent boys. *Nerv. Child*, 1949, 8, 50-53.
2. COOPER, J. M. *Children's Institutions*. Philadelphia: Dolphin, 1931.
3. EDMISTON, R. W., & BAIRD, F. The adjustment of orphanage children. *J. Educ. Psychol.*, 1949, 40, 482-488.
4. FEINBERG, H. Achievement of a group of children in foster homes as revealed by the Stanford Achievement Test. *J. Genet. Psychol.*, 1949, 75, 293-303.
5. FRIED, R., & MAYER, M. F. Socio-emotional factors accounting for growth failure in children living in an institution. *J. Pediat.*, 1948, 33, 444-457.
6. GOLDFARB, W. Variations in adolescent adjustment of institutionally reared children. *Amer. J. Orthopsychiat.*, 1947, 17, 449-451.
7. JASTAK, J. Psychometric personality traits. *Del. State Med. J.*, 1949, 21, 165-176.
8. MARSH, H. R. Psychologist can help in planning for baby's adoption. *Child*, 1949, 14, 68-71.
9. MARTIN, F. A comparative study of the intelligence of children from private homes and of those from an institutional home. *Peabody J. Educ.*, 1947, 24, 198-202.
10. WECHSLER, D. *The Measurement of Adult Intelligence*. Baltimore: Williams & Wilkins, 1941.

191 Joralemon Street
Brooklyn 2, New York

BOOKS

The Journal of Genetic Psychology, the *Journal of General Psychology*, and the *Journal of Social Psychology*, will buy competent reviews at not less than \$2 per printed page and not more than \$3 per printed page, but not more than \$15.00 for a single review.

Conditions. Only those books that are listed below in this section are eligible for such reviews. In general, any book so listed contains one or more of the following traits: (a) Makes an important theoretical contribution; (b) consists largely of original experimental research; (c) has a creative or revolutionary influence in some special field or the entire field of psychology; (d) presents important techniques.

The books are listed approximately in order of receipt, and cover a period of not more than three years. A reviewer must possess the Ph.D. degree or its equal in training and experience.

Procedure. If among the books listed below there is one that seems important to you, you are invited to write a review of that book. It is not necessary to make arrangements with the Editor. Just send in your review. It does not matter if the book in question has been reviewed before.

1952

- STONE, C. P., *Ed.* Annual Review of Psychology, Vol. 3. Stanford, Calif.: Annual Reviews, 1952. Pp. 462.
- HOCH, P. H. & ZUBIN, J., *Eds.* Relation of Psychological Tests to Psychiatry. New York: Grune & Stratton, 1952. Pp. 301.
- CATTELL, R. B. Factor Analysis. New York: Harper, 1952. Pp. 462.
- FOERSTER, H. VON, *Ed.* Cybernetics. New York: Macy Foundation, 1952. Pp. 240.
- BENDER, L. Child Psychiatric Techniques. Springfield, Ill.: 1952. Pp. 335.
- METTLER, F. A., *Ed.* Psychosurgical Problems. New York: Blakiston, 1952. Pp. 357.
- BENDA, C. E. Developmental Disorders of Mentation and Cerebral Palsies. New York: Grune & Stratton, 1952. Pp. 565.
- KINGET, G. M. The Drawing-Completion Test. New York: Grune & Stratton, 1952. Pp. 238.
- WOLFF, C. The Hand in Psychological Diagnosis. New York: Philosophical Library, 1952. Pp. 218.
- WOLFF, W. The Dream: Mirror of Conscience. New York: Grune & Stratton, 1952. Pp. 348.
- GORLOW, L., HOCH, E. L., & TELSCHOW, E. F. The Nature of Nondirective Group Psychotherapy. New York: Teach. Coll., Columbia Univ., 1952. Pp. 143.
- AUSUBEL, D. P. Ego Development and the Personality Disorders. New York: Grune & Stratton, 1952. Pp. 564.
- BYCHOWSKI, G., & DESPERT, J. L. Specialized Techniques in Psychotherapy. New York: Basic Books, 1952. Pp. 371.
- WHITE, R. W. Lives in Progress: A Study of the Natural Growth of Personality. New York: Dryden, 1952. Pp. 376.
- KARN, H. W., & GILMER, B. VON H. Readings in Industrial and Business Psychology. New York: McGraw-Hill, 1952. Pp. 476.
- ROBACK, A. A. History of American Psychology. New York: Library Pub., 1952. Pp. 426.
- KROEBER, A. L. The Nature of Culture. Chicago: Univ. Chicago Press, 1952. Pp. 437.
- FERGUSON, L. W. Personality Measurement. New York: McGraw-Hill, 1952. Pp. 457.

- BECK, S. J. Rorschach's Test: III. Advances in Interpretation. New York: Grune & Stratton, 1952. Pp. 301.
- VINACKE, W. E. The Psychology of Thinking. New York: McGraw-Hill, 1952. Pp. 392.
- KUHLEN, R. G., & THOMPSON, G. G., *Eds.* Psychological Studies of Human Development. New York: Appleton, Century, Crofts, 1952. Pp. 533.
- SNYDER, F. W., & PRONKO, N. H. Vision with Spatial Inversion. Wichita: Univ. Wichita Press, 1952. Pp. 144.
- JERSILD, A. T. In Search of Self. New York: Teach. Coll., Columbia Univ., 1952. Pp. 141.
- HIRSH, O. J. The Measurement of Hearing. New York: McGraw-Hill, 1952. Pp. 364.
- ASHBY, W. R. Design for a Brain. New York: Wiley, 1952. Pp. 260.
- SHAFFER, G. W. Fundamental Concepts in Clinical Psychology. New York: McGraw-Hill, 1952. Pp. 540.
- HAYEK, F. A. The Sensory Order. Chicago: Univ. Chicago Press, 1952. Pp. 209.
- AMES, L. B., LEARNED, J., METRAUX, R. W., & WALKER, R. N. Child Rorschach Responses. New York: Harper, 1952. Pp. 310.

1953

- KLUCKHORN, C., MURRAY, H. A., & SCHNEIDER, D. M., *Eds.* Personality. New York: Knopf, 1953. Pp. 701.
- SKINNER, B. F. Science and Human Behavior. New York: Macmillan, 1953. Pp. 461.
- STONE, C. P., & TAYLOR, D. W., *Eds.* Annual Review of Psychology. Stanford: Annual Reviews, 1953. Pp. 485.
- BURGESS, E. W., & WALLIN, P. Engagement and Marriage. New York: Lippincott, 1953. Pp. 819.
- POWDERMAKER, F. B., & FRANK, J. D. Group Psychotherapy. Cambridge: Harvard Univ. Press, 1953. Pp. 615.
- PHILLIPS, L., & SMITH, J. G. Rorschach Interpretation: Advanced Technique. New York: Grune & Stratton, 1953. Pp. 385.
- MURPHY, G. In the Minds of Men. New York: Basic Books, 1953. Pp. 306.
- BLUM, G. S. Psychoanalytic Theories of Personality. New York: McGraw-Hill, 1953. Pp. 219.
- BROWER, D., & ABT, L. E., *Eds.* Progress in Clinical Psychology. New York: Grune & Stratton, 1953. Pp. 564.
- GELDARD, F. A. The Human Senses. New York: Wiley, 1953. Pp. 365.
- CABOT, H., & KAHL, J. A. Human Relations. Cambridge: Harvard Univ. Press, 1953. Pp. 333 & 273.
- SHERIF, M., & SHERIF, C. W. Groups in Harmony and Tension. New York: Harper, 1953. Pp. 316.
- WHITING, J. W. M., & CHILD, I. L. Child Training and Personality. New Haven, Conn.: Yale Univ. Press, 1953. Pp. 353.

CRITICAL REVIEWS OF RECENT BOOKS

The Journal of Genetic Psychology, 1953, 83, 309-311.

(Humphrey, G. *Thinking*. London: Methuen, 1951. Pp. 331.)

REVIEWED BY WALTER B. PILLSBURY

As would be expected from the author, this is an exhaustive statement of the results of experimental studies of thinking with critical comments. The Würzburg school is taken first, as the step beyond explanation from association. The author accepts imageless thought as a content and appreciates the importance of the task or directive tendency, although he shows skepticism at times as to how it acts. Directive tendency is related to motive. A point is raised as to whether it selects associates as Külpe and his group assume or acts directly as Selz asserted.

For the Gestalt school isolated elements are replaced by larger wholes and directive tendency by the total context of the moment. The school unites in asserting that thinking is productive not reproductive and consists in large part in restructuring the data presented by the problem. Thinking arises from inner tensions.

The last three chapters are devoted to aspects or accompaniments of thinking. A full summary is given of the work on the relation between thinking and movement. Watson and later Jacobsen asserted a close connection, if not an identity, between thinking and movement. Movement accompanies all thinking, as general tension that begins as the problem presents itself, and more specific movements are appropriate to definite problems. The author denies that movement is the complete explanation of thinking, in fact movement may inhibit thought.

The same is said of language which is originally a form of movement. Thinking is not always in words, certainly words are not identical with thought. Speech is an immediate motor response to a situation. No image need precede the series of words, so it is not safe to say that words express thought. It is convenient, but it is difficult to say what is expressed.

Generalization is discussed last instead of first as some have done. Experimental results with three different methods are outlined from Külpe to Heidbreder. These are the Hull method of attaching nonsense syllables to artificial percepts; the multiple choice method of attaching concepts selected by the experimenter to groups of perceptual units; and the language method

of attaching a superordinate to series of particular words. The tests show that no set images or structures characterize the concept, although images are nearly always present, especially early in the development of a concept. The general may arise from the accentuation of some phase or part of the experience; it may appear from assimilation with a familiar object or experience, it may be helped by voluntary intent. Words and images both are present, but may be helpful or harmful in different circumstances. If either is too definite or fixed, generalization is hindered. On the other hand vagueness and generality are not to be confused.

Humphrey questions or denies the importance of *Aufgabe* or set and is very critical of meaning. The last sentence in the book is: "An artificial problem of meaning has been created by treating the *image* and *Speech-activity* apart from their total context." To both of these problems his own description of generalization may well be applied. Psychology has as its problem finding unities of structure and laws of connection in the originally continuous flow of experience. The early men noted the repetition of connections once made and named it association. In the last half century men have been struck by variation in perception and in recall as contexts change. In 1893-4 I noticed that expectation changed readings given to words. The Würzburg school a few years later were struck by the influence of task upon the recall of experience and of the setting in problem solving. The fact was given different names and slightly different functions from Külpe through Ach to Selz, but can all be regarded as the same generalization or concept. Maier's *direction* is probably to be brought under the same head, although it has a different emphasis and is more passive. Humphrey's doubt may be due to lack of picturing the method of action of a setting and fear of developing a faculty. To the first objection we may say that no analysis of any psychological control has been given. The same may be said of physical laws. No definite picture is given for gravitation or the expansion of a metal with heat. All are generalizations from experience that are given names and come to be thought of as reals. They are convenient but mislead if regarded as physical existences.

Meaning is in much the same position. Since Brentano, Bradley, and Bosanquet, it has been apparent a unit of experience is more important for what it represents than for what it is. Since sensation had been set up on the analogy of the atom as the unit of content, it was natural to add meaning to sensation as a different element instead of treating it as a phase of presentation. It is entirely logical to assert that both meaning and sensation should be retained, but meaning is still a psychological problem. Both are useful

concepts. Meaning certainly exists as function if not as structure.

Humphrey also remarks incidentally upon the importance of accepting or rejecting the products of thinking. This corresponds to recognition in recall. I have called it *belief* in thinking, although belief is used somewhat differently in other contexts. Recognition and belief are more accurate than the processes they censor and persist longer in aging and in pathological conditions. Lund has experimented upon *belief* used in this sense. Probably the belief process is what really concerns formal logic rather than problem solving, although it assumes to deal with the latter.

The book gives a valuable summary of a complicated field. The critique is cogent. The reviewer's comments present another view that may be considered in the final estimate of the controversial discipline.

Department of Psychology
University of Michigan
Ann Arbor, Michigan

BOOKS RECENTLY RECEIVED

(There will always be two pages of book titles, listed in the order of receipt, i.e., the most recently received books will be found at the end of the list.)

- TRENAMAN, J. *Out of Step*. New York: Philosophical Library, 1952. Pp. 217.
- HOYLES, J. A. *The Treatment of the Young Delinquent*. New York: Philosophical Library, 1952. Pp. 274.
- RUNES, D. D. *The Societ Impact on Society*. New York: Philosophical Library, 1953. Pp. 202.
- EASTON, D. *The Political System*. New York: Knopf, 1953. Pp. 320.
- HOCH, P. H., & ZUBIN, J. *Current Problems in Psychiatric Diagnosis*. New York: Grune & Stratton, 1953. Pp. 291.
- WHITING, J. W. M., & CHILD, I. L. *Child Training and Personality*. New Haven: Yale Univ. Press, 1953. Pp. 353.
- ANES, L. B., LEARNED, J., NETRAUX, R. W., & WALKER, R. N. *Child Rorschach Responses*. New York: Harper, 1952. Pp. 310.
- ANDERSON, V. A. *Improving the Child's Speech*. New York: Oxford Univ. Press, 1953. Pp. 333.
- HULETT, J. E., & STAGNER, R. *Problems in Social Psychology*. Urbana: Univ. Illinois Press, 1953. Pp. 271.
- KARPF, F. B. *The Psychology and Psychotherapy of Otto Rank*. New York: Philosophical Library, 1953. Pp. 129.
- PODOLSKY, E., *Ed.* *Encyclopedia of Aberrations*. New York: Philosophical Library, 1953. Pp. 550.
- SIMPSON, G. E., & YINGER, J. M. *Racial and Cultural Minorities*. New York: Harper, 1953. Pp. 773.
- BARKER, R. G., *Ed.* *Adjustment to Physical Handicap and Illness*. New York: Social Science Research Council, 1953. Pp. 440.
- SCHULER, E. A., MOWITZ, R. J., & MAYER, A. J. *Medical Public Relations*. Ann Arbor, Mich.: Edwards, 1952. Pp. 216.
- WHITAKER, C. A., & MALONE, T. P. *The Roots of Psychotherapy*. New York: Blakiston, 1953. Pp. 236.
- JUNG, C. G. *Psychology and Alchemy*. (Trans. by R. F. C. Hull.) New York: Pantheon, 1953. Pp. 553.
- OVERHOLSER, W. *The Psychiatrist and the Law*. New York: Harcourt, Brace, 1953. Pp. 147.
- CRUZE, W. W. *Adolescent Psychology and Development*. New York: Ronald Press, 1953. Pp. 557.
- SHERIF, M., & SHERIF, C. W. *Groups in Harmony and Tension*. New York: Harper, 1953. Pp. 316.
- WOOLF, M. D., & WOOLF, J. A. *The Student Personnel Program*. New York: McGraw-Hill, 1953. Pp. 416.
- TIPTON, J. H. *Community in Crisis*. New York: Teachers Coll., 1953. Pp. 180.
- MACE, G. A., & VERNON, P. *Current Trends in British Psychology*. London: Methuen, 1953. Pp. 262.
- TURNER, M. L. *Ships Without Sails*. London: Univ. London Press, 1953. Pp. 160.
- LANGFELD, H. S., BORING, E. G., WERNER, H., & YERKES, R. N., *Eds.* *A History of Psychology in Autobiography*, Vol. IV. Worcester, Mass.: Clark Univ. Press, 1953. Pp. 356.
- TUCKMAN, J., & LORGE, I. *Retirement and the Industrial Worker*. Teach. Coll., Columbia Univ., 1953. Pp. 105.
- BONNER, H. *Social Psychology*. New York: American Book, 1953. Pp. 439.
- COLE, L. E. *Human Behavior*. New York: World Book, 1953. Pp. 884.
- SAYLES, L. R., & STRAUSS, G. *The Local Union*. New York: Harper, 1953. Pp. 269.

- BAKER, H. J. *Introduction to Exception Children*. New York: Macmillan, 1953. Pp. 500.
- PATTERSON, C. H. *The Wechsler-Bellevue Scales*. Springfield, Ill.: Thomas, 1953. Pp. 146.
- BURGER, M. *Einführung in die Pathologische Physiologie*. Leipzig: Thieme, 1953. Pp. 161.
- CABOT, H., & KAHL, J. A. *Human Relations*. (2 vols.) Cambridge: Harvard Univ. Press, 1953. Pp. 333 & 273.
- NAFTALIN, A., NELSON, B. N., SIBLEY, M. Q., & CALHOUN, D. C. *An Introduction to Social Science*. New York: Lippincott, 1953. Pp. 394, 363, & 372.
- GELDARD, F. A. *The Human Senses*. New York: Wiley, 1953. Pp. 365.
- GAMBS, J. S. *Man, Money, and Goods*. New York: Columbia Univ. Press, 1953. Pp. 339.
- SENN, M. J. E., Ed. *Problems of Infancy and Childhood*. New York: Macy, 1953. Pp. 160.
- WARNER, W. L. *American Life: Dreams and Reality*. Chicago: Univ. Chicago Press, 1953. Pp. 268.
- WOOD, M. H. *Paths of Loneliness*. New York: Columbia Univ. Press, 1953. Pp. 250.
- HALMOS, P. *Solitude and Privacy*. New York: Philosophical Library, 1953. Pp. 181.
- BROWER, D., ABT, L. E., Eds. *Progress in Clinical Psychology*. New York: Grune & Stratton, 1952. Pp. 564.
- BERNHARDT, K. S. *Practical Psychology*. New York: McGraw-Hill, 1953. Pp. 337.
- ASHER, E. J., TIFFIN, J., & KNIGHT, F. B. *Introduction to General Psychology*. Boston: Heath. Pp. 515.
- LOEHRICH, R. *Oneirics and Psychosomatics*. McHenry, Ill.: Compass Press, 1953. Pp. 157.
- MONTAGUE, A. *The Natural Superiority of Women*. New York: Macmillan, 1953. Pp. 205.
- HILGARD, E. R. *Introduction to Psychology*. New York: Harcourt, Brace, 1953. Pp. 659.
- ROSEN, J. N. *Direct Analysis*. New York: Grune & Stratton, 1953. Pp. 184.
- KENT, D. P. *The Refugee Intellectual*. New York: Columbia Univ. Press, 1953. Pp. 317.
- BLUM, G. S. *Psychoanalytic Theories of Personality*. New York: McGraw-Hill, 1953. Pp. 219.
- GELDENHUYS, J. N. *The Intimate Life*. New York: Philosophical Library, 1952. Pp. 96.
- LACEY, O. L. *Statistical Methods in Experimentation*. New York: Macmillan, 1953. Pp. 249.
- VERNON, P. E. *Personality Tests and Assessments*. London: Methuen, 1953. Pp. 220.
- MURPHY, G. *In the Minds of Men*. New York: Basic Books, 1953. Pp. 306.
- VEDDER, C. B., KOENIG, S., & CLARK, R. E., Eds. *Criminology: A Book of Readings*. New York: Dryden, 1953. Pp. 714.
- RIESE, W. *The Conception of Disease: Its History, Its Versions, and Its Nature*. New York: Philosophical Library, 1953. Pp. 120.
- SABINE, P. E. *Atoms, Men, and God*. New York: Philosophical Library, 1953. Pp. 226.
- STRATTON, G. M. *Man: Creator or Destroyer*. London: Allen & Unwin, 1952. Pp. 170.
- PHILLIPS, L., & SMITH, J. G. *Rorschach Interpretation: Advanced Technique*. New York: Grune & Stratton, 1953. Pp. 385.
- POWDERMAKER, F. B., & FRANK, J. D. *Group Psychotherapy*. Cambridge: Harvard Univ. Press, 1953. Pp. 615.
- BURGESS, E. W., & WALLIN, P. *Engagement and Marriage*. New York: Lippincott, 1953. Pp. 819.